remote sensing

MDPI

Article

Spectral Emissivity (SE) Measurement Uncertainties across 2.5–14 µm Derived from a Round-Robin Study Made across International Laboratories

Mary F. Langsdale [1,*], Martin Wooster [1], Jeremy J. Harrison [2,3], Michael Koehl [4], Christoph Hecker [5], Simon J. Hook [6], Elsa Abbott [6], William R. Johnson [6], Alessandro Maturilli [7], Laurent Poutier [8], Ian C. Lau [9] and Franz Brucker [10]

1 NERC National Centre for Earth Observation (NCEO), c/o Department of Geography, King's College London, London WC2B 4BG, UK; martin.wooster@kcl.ac.uk
2 National Centre for Earth Observation (NCEO), Leicester LE1 7RH, UK; jh592@leicester.ac.uk
3 Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH, UK
4 Optosol GmbH, 83614 Miesbach, Germany; michael.koehl@ise.fraunhofer.de
5 Department of Earth Systems Analysis, University of Twente (UT-ITC), Hengelosestraat 99, P.O. Box 37, 7500 AA Enschede, The Netherlands; c.a.hecker@utwente.nl
6 National Aeronautics and Space Administration, Jet Propulsion Laboratory (NASA-JPL) 4800 Oak Grove Drive, Pasadena, CA 91109, USA; simon.j.hook@jpl.nasa.gov (S.J.H.); elsa.a.abbott@jpl.nasa.gov (E.A.); william.r.johnson@jpl.nasa.gov (W.R.J.)
7 Planetary Spectroscopy Laboratory (PSL) at the German Aerospace Centre (DLR), Rutherfordstrasse 2, 12489 Berlin, Germany; alessandro.maturilli@dlr.de
8 ONERA—The French Aerospace Lab, 6 Chemin de la Vauve aux Granges, 91120 Palaiseau, France; laurent.poutier@onera.fr
9 Commonwealth Scientific and Industrial Research Organisation (CSIRO), ARRC Building, 26 Dick Perry Avenue Kensington Western Australia 6151, P.O. Box 1130 Bentley, Kensington 6102, Australia; ian.lau@csiro.au
10 Fraunhofer Institute for Solar Energy System (ISE), Heidenhofstr. 2, 79110 Freiburg, Germany; brucker@ise.fhg.de
* Correspondence: mary.langsdale@kcl.ac.uk

Abstract: Information on spectral emissivity (SE) is vital when retrieving and evaluating land surface temperature (LST) estimates from remotely sensed observations. SE measurements often come from spectral libraries based upon laboratory spectroscopic measurements, with uncertainties typically derived from repeated measurements. To go further, we organised a "round-robin" inter-comparison exercise involving SE measurements of three samples collected at seven different international laboratories. The samples were distilled water, which has a uniformly high spectral emissivity, and two artificial samples (aluminium and gold sheets laminated in polyethylene), with variable emissivities and largely specular and Lambertian characteristics. Large differences were observed between some measurements, with standard deviations over 2.5–14 µm of 0.092, 0.054 and 0.028 emissivity units (15.98%, 7.56% and 2.92%) for the laminated aluminium sheet, laminated gold sheet and distilled water respectively. Wavelength shifts of up to 0.09 µm were evident between spectra from different laboratories for the specular sample, attributed to system design interacting with the angular behaviour of emissivity. We quantified the impact of these SE differences on satellite LST estimation and found that emissivity differences resulted in LSTs differing by at least 3.5 K for each artificial sample and by more than 2.5 K for the distilled water. Our findings suggest that variations between SE measurements derived via laboratory setups may be larger than previously assumed and provide a greater contribution to LST uncertainty than thought. The study highlights the need for the infrared spectroscopy community to work towards standardized and interlaboratory comparable results.

Keywords: FTIR; infrared spectroscopy; directional hemispherical reflectance; spectral emissivity; LWIR; land surface temperature

Citation: Langsdale, M.F.; Wooster, M.; Harrison, J.J.; Koehl, M.; Hecker, C.; Hook, S.J.; Abbott, E.; Johnson, W.R.; Maturilli, A.; Poutier, L.; et al. Spectral Emissivity (SE) Measurement Uncertainties across 2.5–14 µm Derived from a Round-Robin Study Made across International Laboratories. *Remote Sens.* **2021**, *13*, 102. https://doi.org/10.3390/rs13010102

Received: 31 October 2020
Accepted: 21 December 2020
Published: 30 December 2020

1. Introduction

Spectral emissivity (SE) is an intrinsic material property, defined as the ratio (0–1) of the electromagnetic radiation emitted by an object at a particular wavelength to that emitted at the same wavelength by a perfect blackbody at the same thermodynamic temperature [1]. SE values range from 0 to 1 emissivity units, albeit when averaged across the mid-wave infrared (MWIR) or long-wave infrared (LWIR) atmospheric windows most natural materials have an SE higher than 0.4 and 0.6 respectively. SE information is essential for the derivation of land surface temperature (LST), an Essential Climate Variable (ECV) important to the understanding and modelling of many Earth system processes from local to global scales [2,3].

LST retrieval algorithms primarily use remotely sensed observations of electromagnetic radiation in the long-wave infrared (LWIR; 8–14 μm) part of the thermal infrared atmospheric window. Some LST algorithms do make use of the mid-wave infrared (MWIR; 3–5 μm) atmospheric window, though these are less common because daytime MWIR measurements are a mixture of thermally emitted and solar reflected radiation [3]. Usually the specific emissivity information required for use in LST retrieval algorithms is the SE integrated over the spectral response function of each of the spectral measurement channels considered in the algorithm [4], though for convenience we still refer herein to SE since typically it is by knowing this that the band-integrated spectral integrated emissivity values are determined.

However, errors in emissivity typically result in significant LST biases. For example, for typical Earth surface conditions, SE uncertainties of 0.01 deliver typical uncertainties of around 0.6 K in the retrieved LST [5]. Given this and recent experimental studies into angular and structural emissivity dependence [6,7], accurate knowledge of SE has been identified as one of the greatest challenges to retrieving sufficiently precise LST to support a wider range of applications [8].

Remotely sensed LST algorithms typically require either (i) knowledge of the SE or its spectral integral in advance, as with widely used split-window algorithms (for example [9]) or (ii) estimate SE or its spectral integral as part of the retrieval process, as with the Temperature Emissivity Separation (TES) algorithm [10]. Laboratory measurements of SE, typically made using Fourier transform infrared (FTIR) spectrometer setups, are commonly used in both approaches, either when deriving the split-window coefficients [11], for calibration of satellite or airborne sensors [12] or for ground-truthing of the LST and SE outputs [13–15].

Interest in SE measurements has increased in recent years, largely due to advances in thermal remote sensing and a concerted effort to reduce LST retrieval uncertainty following the classification of LST as an ECV. Campaigns such as Fiducial Reference Measurements for validation of Surface Temperature from Satellites (FRM4STS) have focused on such efforts [16,17]. Laboratory measurements of SE are generally considered to be the "truth" in such work, either for measurement of samples that can be transported without modifying the sample and its emissivity or for evaluating the accuracy of field methodologies on appropriate samples [18]. One key advantage of laboratory SE measurements is the highly controlled conditions under which measurements can be made compared to typical field measurement conditions, and potentially the higher spectral resolution often possible with laboratory setups [19]. Laboratory SE measurements are therefore commonly used as reference measurements when comparing them to those derived using field, airborne or satellite observations [20–22].

Given the importance of SE information, multiple laboratories have now developed capabilities for determining SE from thermally emitted or reflected infrared radiation measurements of target samples. Whilst the former "thermal emission" approach is used in for example the SLUM (Spectral Library of impervious Urban Materials) library of Kotthaus et al. [23], the method requires that the samples be heated to well above room temperature, which for some materials is not always possible and which can introduce issues with regards to sample temperature homogeneity when the sample is removed

from the heat source to be measured. The latter "reflected radiation" approach instead illuminates a room temperature sample with infrared radiation and measures how much of the radiation is reflected, with SE then determined through use of Kirchhoff's Law [24]. Key advantages of this approach are that no artificial heating of the samples is required, so all types of sample are analysable, and sample temperature inhomogeneity is not an issue. The approach has been widely applied to provide much of the SE data populating the most commonly used online spectral emissivity libraries, such as the ECOSTRESS (ECOsystem Spaceborne Thermal Radiometer Experiment on a Space Station)—formerly ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer)—spectral library [25]. SE data from the ECOSTRESS spectral library are used within the TES algorithm to provide the spectra required to derive some of the algorithm coefficients [26].

However, the quality of laboratory SE measurements is not always apparent, and there are relatively few reflectance standards readily available for use in the MWIR and LWIR spectral regions with which to assess this, unlike in the near-infrared (NIR) or shortwave infrared (SWIR) [27]. SE quality metrics for an individual laboratory's SE measurements have often been provided as uncertainty values based on repeated measurements of the sample with the same equipment (for example [28]), but comparisons of laboratory SE measurements derived for the same samples but with different equipment and laboratory setups are rare [27,29]. Here we redressed this gap through a "Round-Robin" study involving seven international laboratories all measuring the same set of reference samples whose SE they determined using their own equipment and measurement protocols. The differing SE measurements were intercompared and their inconsistencies explored to understand the impact that any identified differences in SE would have on remotely sensed LST determination.

The lead investigators of this "Round-Robin" study are based at the National Centre for Earth Observation (NCEO) in King's College London (KCL), and their SE measurement setup shown in Figure 1 uses a very similar set of equipment to that used in the Department of Earth Systems Analysis at the University of Twente (UT-ITC) and detailed in [27]. SE measurements of the target sample are inferred from reflected infrared radiation measurements made by a Bruker VERTEX 70 spectrometer and application of Kirchhoff's Law, with the sample positioned under a port of a diffuse highly reflective gold-coated integrating sphere and illuminated by intense radiation coming from an external mid-infrared (MIR) source (Figure 1). Hecker et al. [27] describe two measurement approaches to derive SE from these types of reflectance measurements—the substitution and comparative calibration methods that are described in detail below.

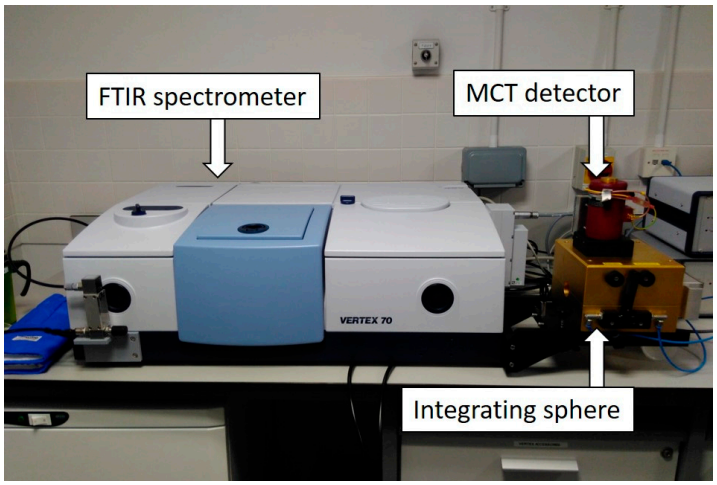

Figure 1. Laboratory setup for surface spectral emissivity determination at King's College London (KCL), based on a Bruker VERTEX 70 FTIR spectrometer, an external source of high intensity thermal radiation, and a gold-coated integrating sphere with a MCT (mercury-cadmium-telluride) detector cooled with liquid nitrogen.

The substitution method of SE derivation uses a material of known emissivity (in this case a Labsphere Infragold™ plate) as a reference sample, and this is first placed under the sample port (e.g., of the integrating sphere of Figure 1) and a "reference measurement" made. The reference sample is then replaced by the target sample and the "sample measurement" made. Spectral reflectance is then calculated through the ratio of these two measurements. In the comparative calibration method of SE derivation, the sample and reference samples are mounted simultaneously (e.g., through multiple ports or use of the internal sphere wall as for the reference), their measurements made consecutively (e.g., through use of an internal rotating mirror), with the sample spectral reflectance again calculated through their ratio. Theoretically, the comparative method should provide some benefit since it avoids a known limitation in the substitution method (the so-called "substitution error"), where changes in the total internal sphere reflectance between measurements of the reference and the sample cause underestimation of sample reflectances (and thus overestimation of emissivity) as discussed in [30]. Hardy and Pineo [31] determined that the substitution error could be as much as 25% for low reflectance samples and 12% for samples with medium reflectance. Corrections have been developed [32,33] but even these are known to include errors of up to 1% from approximations in the calculations. However using both the substitution and comparative methods with the setup at UT-ITC, Hecker et al. [27] observed differences between the SEs derived, with those calculated using the substitution method in closer agreement with other spectra, thus questioning the assumption that the comparative method provided improved results. They attributed these differences to variations in the measurement geometries between the reference and sample measurements made using the comparative method. The measurement setup at the KCL laboratory has been designed to attempt to overcome this issue, and design specifications were to have as identical a path length as possible between the sample and reference measurement when using the comparative method. Within the current work we will therefore also assess the relative performance of the KCL setup when performing SE retrieval with these two different measurement approaches.

Our objectives are therefore threefold: (i) investigate the consistency of SE measurements derived from measurements made in different international laboratories through a Round Robin intercomparison study using reference samples, (ii) evaluate the substitution and comparative calibration methods of SE measurement using the setup shown in Figure 1, and (iii) assess the impact that any SE differences and uncertainties stemming from the results of (i) and (ii) have on typical satellite LST retrievals.

2. Materials and Methods

2.1. SE Sample Standards

Two artificial SE samples were used in this study, with one specular and one diffuse to test the setup performances for samples with different scattering properties. Sample 1 was a thin 60 mm × 60 mm aluminium sheet showing specular reflective behaviour and laminated in polyethylene (PE). Sample 2 was a thin 47 mm × 57 mm diffusely reflective gold sheet also laminated in PE. The samples are shown in Figure 2 and were selected primarily for their robustness and their ability to be used in multiple different laboratory setups having different sample holder sizes and different measurement alignments (e.g., side-looking and down-looking instrumentation). Additionally, both have known spectrally varying properties over the 2.5–14 μm spectral region since the metal foils have low emissivity but the PE film has spectral regions of high absorptivity that result in spectral regions of high emissivity [34]. Due to the nature of the materials used in these reference samples, measurements of SE that required heating of the samples are inappropriate and only laboratories where SE measurements are obtained from directional hemispherical reflectance (DHR) setups were considered.

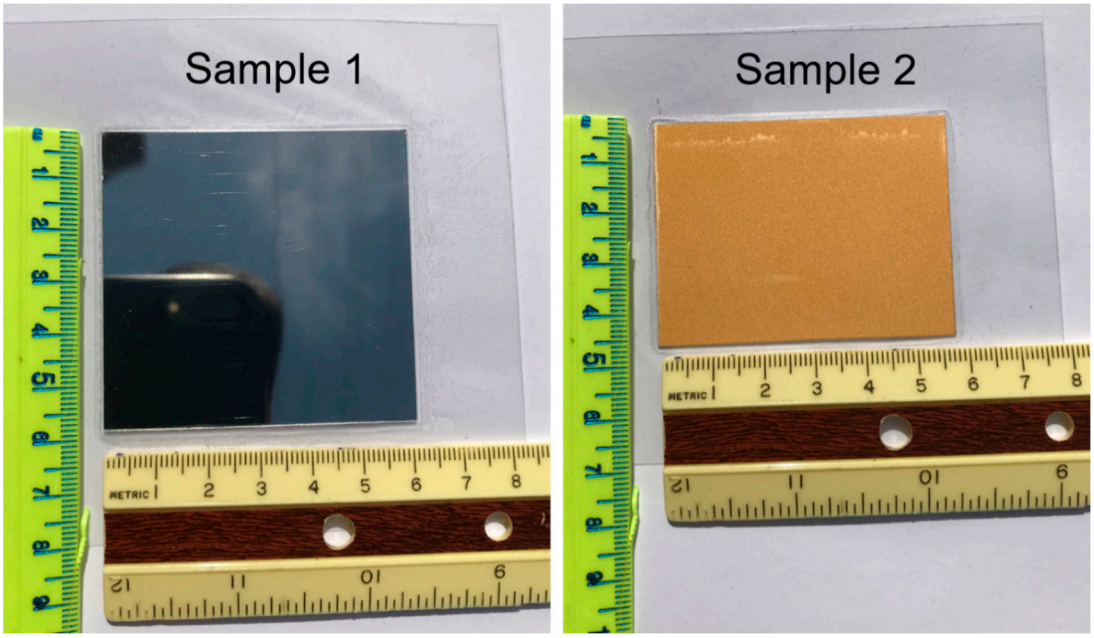

Figure 2. Artificial samples used within the Round Robin study of surface spectral emissivity described herein. Sample 1 was a 60 mm × 60 mm aluminium sheet laminated in a polyethylene covering to deliver a more specularly reflecting sample, whilst Sample 2 was a 47 mm × 57 mm gold sheet laminated in a polyethylene covering to deliver a more Lambertian reflecting sample. Both samples showed spectral regions of both high and low emissivity. A third SE reference sample consisting of distilled water was also used within laboratories capable of measuring liquid samples.

In addition to the two artificial SE standards shown in Figure 2, the laboratories participating in the Round Robin were also requested to make measurements of a sample of distilled water if their setup was permitting. Distilled water is widely available and its SE is well known and available in the ECOSTRESS spectral library, with SE measurements of distilled water featuring in other inter-comparison studies including [27]. Additionally, since distilled water has a low spectral reflectance of only a few percent, the retrieval of accurate SE poses a challenge for the signal-to-noise capability of setups that measure in DHR mode [28], thus providing a good test of a laboratory's capability.

2.2. Laboratory SE Measurement Setups and Schedule

In total twelve different measurement setups conducted at seven different international research organisations were used to measure the SE of the two artificial sample standards. All setups other than the Agilent 4300 Handheld FTIR at CSIRO (Commonwealth Scientific and Industrial Research Organisation) involved integrating spheres to derive spectrally resolved hemispherical reflectance (ρ) measurements over the 2.5–14 µm spectral region. The Agilent does not have an integrating sphere but instead uses mirrors to capture diffuse reflectance from a target. Five of the laboratories provided additional SE measurements of distilled water. Included in the distilled water comparison is a distilled water spectrum measured by John Hopkins University from the ECOSTRESS spectral library [25], also considered in [27] and which was found by [28] to be within 0.17% of a theoretical spectrum of distilled water calculated using optical constants of water in the infrared [35]. This spectrum is identifiable as ECOSTRESS Spectral Library (ESL).

The setup and instruments used within each of the seven laboratories vary in design, interferometer type, age, reference standard (typically gold), and general measurement protocols. A detailed description of each setup is provided in Appendices A–G, with an overview in Table 1. Note that CSIRO were able to use two different instruments in the Round Robin—a Bruker VERTEX 80v with an integrating sphere with multiple possibilities of port placement for the sample and reference standards, and an Agilent 4300 Handheld FTIR.

Table 1. An overview of the spectrometer setups of the multiple research organisations participating in this Round Robin surface spectral emissivity measurement intercomparison. The setups are presented here in the order they were used to measure the SE of the selected samples, with information about the instrument, reference standard, gases purging from and references in the literature where available. The names in brackets identify how the laboratories will be referred to subsequently.

Research Institution + Country		Instrument Specification	Reference Standard	Purge Gas	References
National Centre for Earth Observation at Department of Geography, King's College London (KCL)	UK	Bruker VERTEX 70 FTIR spectrometer with external gold-coated integrating sphere, external detector and external water-cooled globar infrared source	Internal Infragold sphere wall (Comparative) Infragold Labsphere target (Substitution)	Dry, CO_2 free air	-
Fraunhofer Institute for Solar Energy System ISE, Optosol GmbH (Optosol)	DE	Bruker VERTEX 80 FTIR spectrometer modified with external integrating sphere and detector	-	-	-
Department of Earth Systems Analysis, University of Twente (UT-ITC)	NE	Bruker VERTEX 70v FTIR spectrometer with external gold-coated integrating sphere, external detector and external water-cooled globar infrared source	Infragold standard	Dry, CO_2 free air	[19,27,36]
Planetary Spectroscopy Laboratory at the German Aerospace Center (DLR)	DE	Bruker 80v FTIR spectrometer with gold-coated integrating sphere	Infragold standard	-	[37]
National Aeronautics and Space Administration-Jet Propulsion Laboratory (NASA-JPL)	US	Nicolet 6700 FTIR spectrometer with an external Labsphere integrating sphere	Infragold standard	Dry, CO_2 free air	[38]
ONERA—The French Aerospace Lab (ONERA)	FR	Bruker Equinox 55 FTIR spectrometer equipped with a Labsphere Infragold-coated integrating sphere	Infragold Labsphere target	Dry, CO_2 free air	[39]
Commonwealth Scientific and Industrial Research Organisation (CSIRO)	AUS	(1) Bruker VERTEX 80v spectrometer with Bruker integrating sphere with multiple sample ports (2) Agilent 4300 Handheld FTIR spectrometer	1. Infragold target 2. Coarse silver target	1. Unknown if operational 2. None	[40]

The Round Robin was organised sequentially, in that the artificial reference SE samples were measured by one laboratory and then sent onto the next, in the order shown in Table 1. Samples of distilled water were provided by the laboratory participants themselves from local supplies. Measurements commenced in September 2017 at KCL and Optosol. Samples were returned to KCL both halfway (Oct 2018) and at the end (Sept 2019) of the exercise to be remeasured on the same setup and with the same methodology to check for absolute changes to the samples. ONERA's Infragold reference standard was additionally sent to KCL to be measured using the KCL setup to determine the impact of using a different laboratory's reference standard to derive reflectance.

2.3. Measurement Protocol

To enable easy intercomparison of the final SE data, all participating laboratories were requested to make spectral reflectance measurements across the 2.5–14 μm spectral range (4000–714 cm^{-1}) at spectral resolutions of 4 cm^{-1}, 8 cm^{-1} and (if capable) 0.5 cm^{-1}, with all other settings kept as their standard protocol. The majority of participant laboratories made their sample reflectance measurements (ρ_s) by comparing the measured radiation reflected from the sample (V_s) with that reflected from a reference standard (V_r) of known (and near unity) spectral reflectance across the wavelength range of interest such that:

$$\rho_s(\lambda) = \frac{V_s(\lambda)}{V_r(\lambda)}\rho_r(\lambda) \quad (1)$$

where λ is wavelength and $\rho_r(\lambda)$ the spectral reflectance of the reference standard, generally provided by manufacturer or from previous calibration. KCL and UT-ITC made additional open port measurements and subtracted these from both the sample and reference measurements during the spectral reflectance calculation in Equation (1) in order to remove background radiation, as detailed in [27]. ONERA, KCL and NASA-JPL made additional adjustments to the measured data to compensate for the substitution error described in Section 1.

Sample emissivity (ε_s) was inferred from measured reflectance using Kirchhoff's law [24]:

$$\varepsilon_s(\lambda) = 1 - \rho_s(\lambda) \quad (2)$$

The number of scans varied between the different methods, ranging from 60 scans for each sample at CSIRO using the Agilent 4300 Handheld FTIR to eight repeat measurements of 512 scans for each sample at UT-ITC. For most setups, a full measurement sequence (reference, sample and any additional measurements) took between 5 and 30 min.

A summary of the method of SE determination used by each participant laboratory is presented in Table 2, with abbreviations identifying how the individual measurements will be subsequently referred to. Detailed descriptions of the individual measurement protocols are presented in Appendices A–G.

Table 2. Details of the methods of spectral emissivity (SE) determination used by each participant laboratory, with abbreviations identifying how the individual measurements will be subsequently referred to.

Institute	Measurement Description	Measurement Abbreviation
KCL	(1) Substitution method with sample and reference alternately placed under bottom port; sphere correction and open port subtraction applied	KCL_sub
	(2) Comparative method with sample under bottom port and internal gold wall of the integrating sphere as reference; open port subtraction applied	KCL_comp
Optosol	Comparative method using internal gold wall of the integrating sphere as reference	Optosol
NASA-JPL	Substitution method with sample and reference alternately placed under bottom port; sphere correction applied	NASA-JPL
UT-ITC	Substitution method with sample and reference alternately placed under bottom port; open port subtraction applied	UT-ITC
DLR	Substitution method with sample and reference alternately placed over sample port	DLR
ONERA	Substitution method with sample and reference alternately placed under sample port. A second measurement of each is made with a tilted beam to correct for substitution method error	ONERA

Table 2. *Cont.*

Institute	Measurement Description	Measurement Abbreviation
CSIRO	(1) Comparative method with reference under the bottom port of the Bruker VERTEX 80v FTIR spectrometer and sample in the top port	CSIRO_BG-B_S-T
	(2) Comparative method with reference in the top port of the Bruker VERTEX 80v FTIR spectrometer and sample under the bottom port	CSIRO_BG-T_S-B
	(3) Comparative method using the internal wall of the integrating sphere of the Bruker VERTEX 80v FTIR spectrometer as reference and sample under the bottom port	CSIRO_BG-W_S-B
	(4) Comparative method using the internal wall of the integrating sphere of the Bruker VERTEX 80v FTIR spectrometer as reference and sample in the top port	CSIRO_BG-W_S-T
	(5) Substitution method for the Agilent 4300 Handheld FTIR spectrometer	CSIRO_Agilent

3. Results

Due to the very strong similarities in the results obtained at the different spectral resolutions used, we report here the 4 cm^{-1} results only. This resolution was used by all laboratories and the one for which the widest intercomparisons can be made. Conclusions from the other spectral resolution measurements were similar.

3.1. Sample Stability

Figure 3 shows the mean and standard deviation derived for the two artificial samples from three measurements collected at KCL (with identical parameters and setup) at the start, midway through and at the end of the exercise as detailed in Section 2.2. Despite slight physical abrasion observed on Sample 1 at the end of the study, differences between the emissivity measurements made at different times are small for both samples, as shown by the generally low standard deviation (< 0.01 for most wavelengths, with the peak in standard deviation around 4.3 μm attributed to insufficient purging during one measurement as it appears in the CO_2 region). Variability is slightly greater for the specular sample (Sample 1) than the diffuse sample (Sample 2), with the mean standard deviations across $2.5 - 14$ μm 0.009 and 0.006 respectively for these two samples. However, these levels of variability are still within the ranges observed in the reproducibility tests in similar studies [27]. These data indicate that the absolute SEs of Samples 1 and 2 were stable throughout the experiment, and that any SE differences found between the different laboratory measurements cannot be attributed to changes in the samples over time.

3.2. Comparison between Different Laboratory's Emissivity Measurements

3.2.1. Absolute Differences between Emissivity Measurements

The SEs of the two artificial samples and distilled water measured using the setups listed in Table 2 are shown in Figures 4 and 5 respectively. From Figure 4, we can see that there are some large SE differences between the measurements of both artificial samples at certain wavelengths. These differences are reduced for distilled water (Figure 5), although there were fewer measurements here as not all laboratories were able to make measurements of this sample. For all three samples however, there appear to be a group of measurements within the LWIR region consisting of some of those from CSIRO with lower emissivities (around 0.07 and 0.05 less than the majority of the measurements for the artificial samples and distilled water respectively). In the case of the artificial samples, there also appear to be a top group with spectra from DLR and NASA-JPL, which are around 0.05 higher than the majority of the measurements, although the distilled water measurements from these two laboratories are in close agreement with most of the others.

Variability is observed to be wavelength dependent and is greater in the MWIR than in the LWIR region for all samples, as shown by the standard deviations for the two regions presented in Table 3. This is likely due to increased atmospheric absorption in the MWIR—where there are strong absorption bands for CO_2 and H_2O compared to in the

LWIR region—and the differences in how each setup compensate for these atmospheric effects (if at all). There seems to have been an issue with the CO_2 purging in the DLR setup when measuring the artificial samples as these results report an increase in emissivity in the CO_2 absorption band (4.3 µm), which is not present in the measurements from the other laboratories as can be seen in Figure 4. This is not apparent in the DLR measurement of distilled water however (Figure 5).

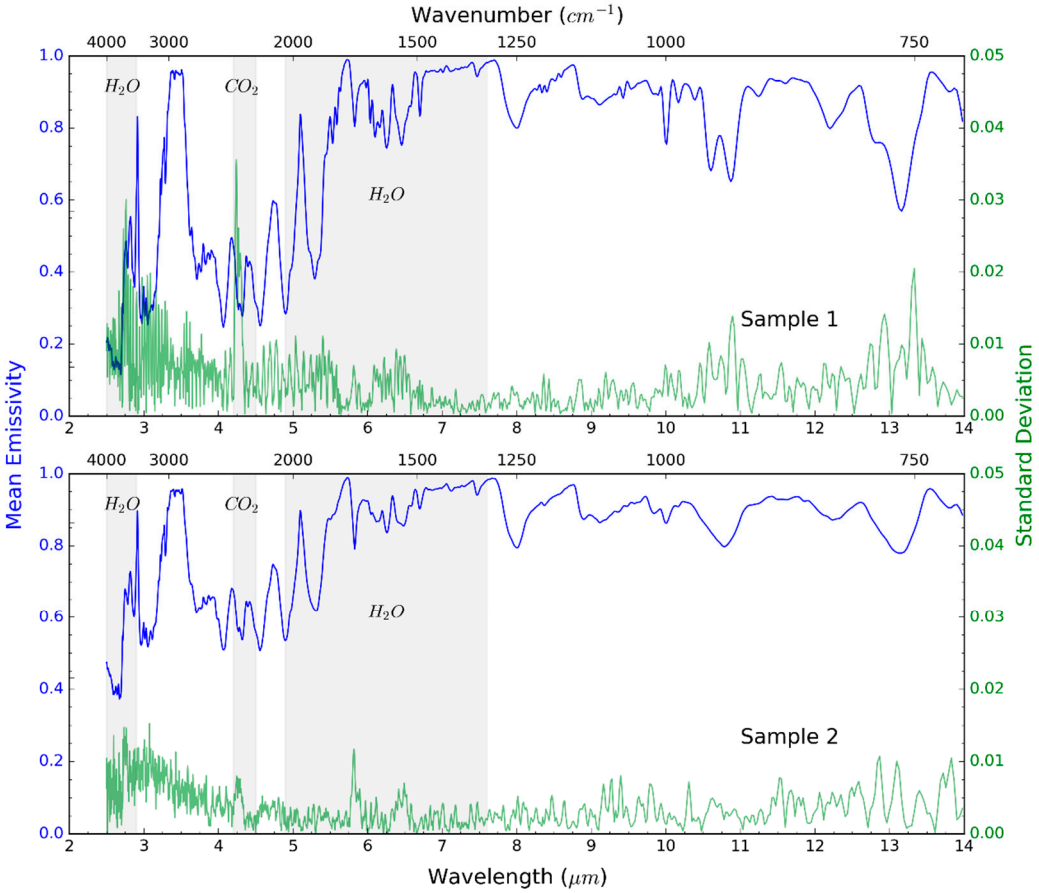

Figure 3. Mean (blue, left axis) and standard deviation (green, right axis) of three spectral emissivity measurements made of Samples 1 and 2 at King's College London (KCL) at the start of the Round Robin (25 September 2017), midway through (9 October 2018), and at the end (11 September 2019). Each sample emissivity was calculated using the substitution method of calibration, using identical measurement setups, parameters and procedures. The absorption bands of relevant gases (H_2O and CO_2) are indicated through the grey vertical bars.

As shown by the standard deviations in Table 3, SE differences were largest for the specular sample (Sample 1), where the standard deviation of the measurements was ±0.092 over the full wavelength range (2.5–14 µm). The maximum observed difference between two measurements (DLR and CSIRO BG-W_S-T, Sample 1) was 0.762 emissivity units, but this occurred around 4.3 µm and was thus within the CO_2 absorption band and seemed likely to be associated with an insufficient atmospheric compensation in the DLR system as discussed earlier. However, DLR and CSIRO consistently produced the highest and lowest emissivities respectively as evident from Table 4, which presents the mean absolute differences of each individual measurement from the mean of all measurements. The measurements made at DLR had the greatest positive bias compared to the mean while all measurements made at CSIRO using the VERTEX 80v FTIR spectrometer (with the exception of that with the reference in the lower port and the sample in the top port, CSIRO BG-B_S-T) had the greatest negative biases compared to the mean.

Table 3. Mean (μ) and standard deviation (σ) of SE (ε) for each sample averaged over a specified wavelength range, calculated using all spectra. The number in brackets shows the standard deviation as a percentage of the mean. The subscripts a, b and c refer to the wavelength ranges averaged over, where a = 2.5–14 μm, b = 3–5 μm (MWIR region) and c = 8–14 μm (LWIR region).

		Spectral Emissivity (μ ± σ)
Sample 1	ε_a	0.574 ± 0.092 (15.98%)
	ε_b	0.476 ± 0.107 (22.54%)
	ε_c	0.855 ± 0.046 (5.33%)
Sample 2	ε_a	0.713 ± 0.054 (7.56%)
	ε_b	0.659 ± 0.060 (9.14%)
	ε_c	0.877 ± 0.037 (4.19%)
Distilled Water	ε_a	0.962 ± 0.028 (2.92%)
	ε_b	0.954 ± 0.030 (3.18%)
	ε_c	0.970 ± 0.024 (2.52%)

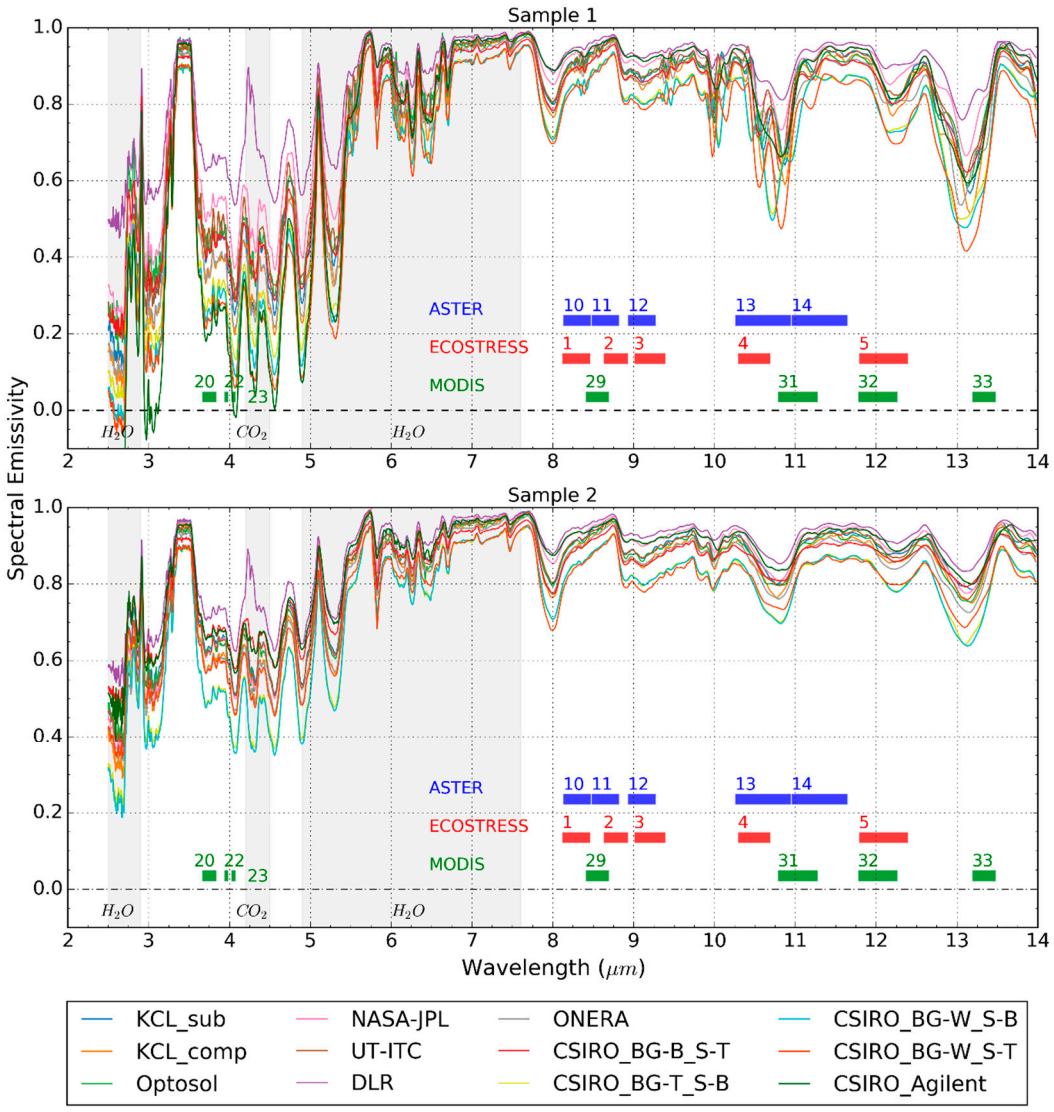

Figure 4. Surface spectral emissivity measurements for artificial Samples 1 and 2, measured at 4 cm^{-1} resolution over the 2.5–14 μm spectral range. The spectral ranges of MODIS bands 20, 22, 23, 29 and 31–33 are indicated, which are used to retrieve LST using the MWIR and LWIR regions in the day/night algorithm detailed in [41]. ECOSTRESS and ASTER thermal band locations are also shown in green, red and blue respectively. The absorption bands of relevant gases (H_2O and CO_2) are indicated through the grey vertical bars. For legend abbreviations see Table 2.

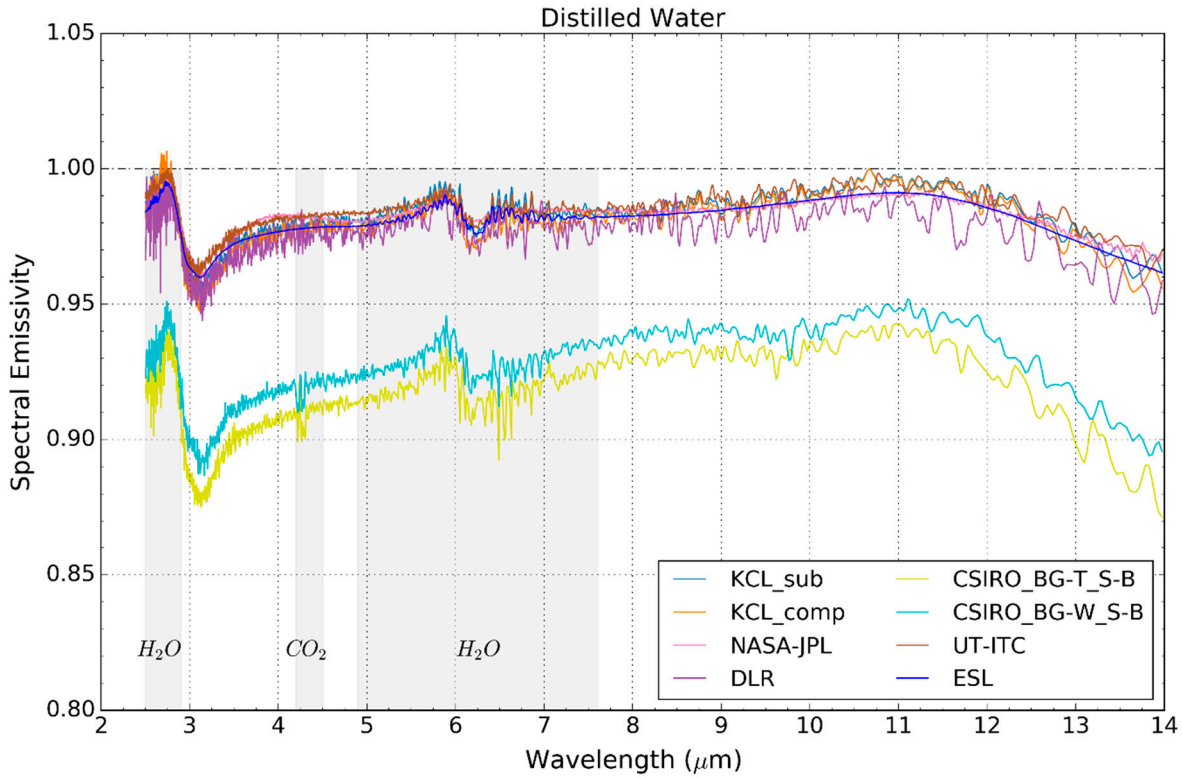

Figure 5. SE measurements of distilled water collected at a 4 cm^{-1} spectral resolution over the 2.5–14 µm wavelength range. Additionally, shown is the distilled water SE spectrum from the ECOSTRESS Spectral Library (ESL) [25]. The absorption bands of relevant gases (H_2O and CO_2) are indicated through the grey vertical bars. As before, for legend abbreviations see Table 2.

Table 4. The mean SE differences for each setup over specified wavelength ranges, calculated as $(\varepsilon – \varepsilon_{mean})$, where the mean emissivity ε_{mean} was calculated using all spectra. As before, the subscripts a, b and c refer to the wavelength ranges averaged over, where a = 2.5–14 µm, b = 3–5 µm (MWIR region) and c = 8–14 µm (LWIR region).

	Sample 1			**Sample 2**			**Distilled Water**		
	$\Delta\varepsilon_a$	$\Delta\varepsilon_b$	$\Delta\varepsilon_c$	$\Delta\varepsilon_a$	$\Delta\varepsilon_b$	$\Delta\varepsilon_c$	$\Delta\varepsilon_a$	$\Delta\varepsilon_b$	$\Delta\varepsilon_c$
CSIRO_Agilent	−0.113	−0.145	0.031	0.052	0.056	0.036	-	-	-
CSIRO_BG-B_S-T	0.030	0.030	0.002	0.040	0.049	−0.001	-	-	-
CSIRO_BG-T_S-B	−0.070	−0.087	−0.057	−0.085	−0.101	−0.055	−0.049	−0.053	−0.043
CSIRO_BG-W_S-B	−0.090	−0.111	−0.061	−0.091	−0.107	−0.056	−0.039	−0.042	−0.033
CSIRO_BG-W_S-T	−0.107	−0.132	−0.073	−0.035	−0.035	−0.057	-	-	-
DLR	0.183	0.229	0.069	0.088	0.105	0.053	0.013	0.015	0.010
KCL_comp	−0.014	−0.014	−0.004	−0.013	−0.014	0.001	0.016	0.016	0.016
KCL_sub	0.023	0.028	0.017	0.024	0.028	0.020	0.019	0.019	0.017
NASA-JPL	0.089	0.113	0.046	0.015	0.014	0.035	0.019	0.022	0.015
ONERA	−0.018	−0.016	−0.006	−0.022	−0.022	−0.005	-	-	-
Optosol	0.050	0.052	0.015	0.017	0.012	0.010	-	-	-
UT-ITC	0.037	0.052	0.021	0.010	0.016	0.018	0.021	0.023	0.018

Fewer participants made SE measurements of distilled water (Figure 5), but there was greater agreement among these than for the artificial samples, with reduced standard deviations in all spectral regions (Table 3). Much of the variation in the distilled water spectra appeared due to noise, given the difficulties of measuring a high emissivity (low reflectance) sample on a DHR setup. The CSIRO measurements of distilled water made using the VERTEX 80v spectrometer have a negative bias (0.05) compared to the other laboratories, which was consistent with the results of both artificial samples from this setup. In contrast, the measurements from the other laboratories (DLR, KCL, NASA-JPL and

UT-ITC) were in very close agreement with the spectrum from John Hopkins University available in the ECOSTRESS spectral library (ESL). This can be observed in Figure 6, where the mean and standard deviation of the differences between the spectra of distilled water from the ECOSTRESS spectral library and the measurements of distilled water from DLR, KCL, NASA-JPL and UT-ITC are shown. The spectrum from NASA-JPL over the LWIR spectral region in particular is in close agreement with that from the ECOSTRESS spectral library, reflecting the reduced noise observed in the NASA-JPL distilled water spectrum compared to the other laboratories (Figure 5).

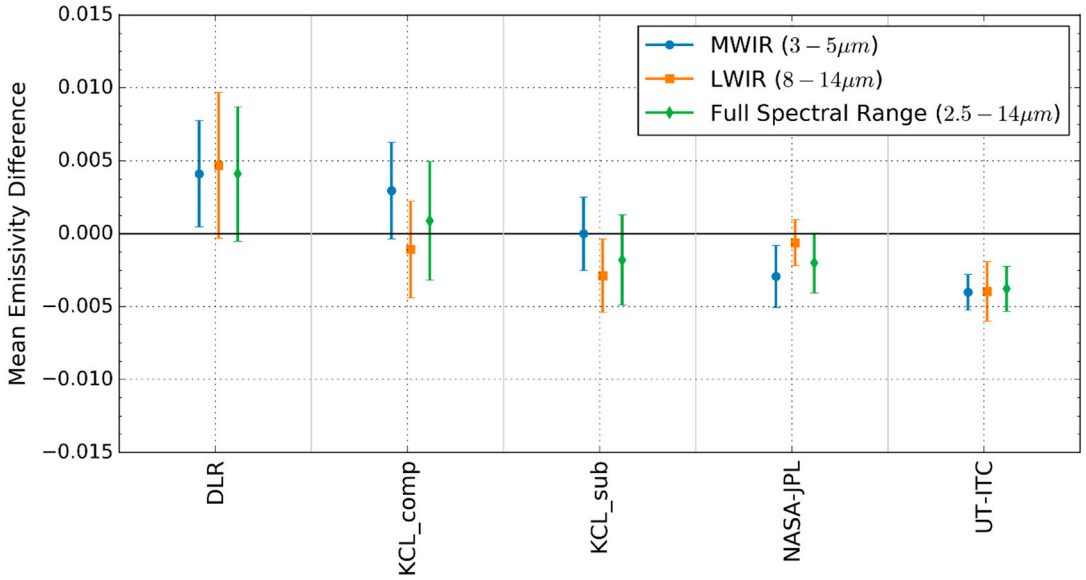

Figure 6. Spectral emissivity measurement differences between the spectra of distilled water from the ECOSTRESS spectral library and all measurements of distilled water aside from those from CSIRO. Differences ($\Delta\varepsilon$) were calculated as $\Delta\varepsilon = \varepsilon_{lab} - \varepsilon_{ESL}$ where ε_{lab} and ε_{ESL} indicate the laboratory measurement and the ECOSTRESS spectral library spectrum respectively, before being averaged over specified spectral ranges. The bars are centred on the mean difference of that spectral range and have a half-length equal to the standard deviation of the difference over that spectral range.

Figures 7 and 8 compare each laboratory's SE measurements against the mean and standard deviation of all SE measurements for Sample 1 and Sample 2 respectively over the LWIR atmospheric window (8–14 µm). In this spectral region, which was most commonly used for remote sensing of LST, the measurements of the artificial samples from DLR and NASA-JPL were consistently above the mean, with the DLR spectra greater than one standard deviation away from the mean and with shallower absorption features at 10.8 µm and 13 µm for Sample 1 (Figure 7). Results from CSIRO and KCL in this same spectral region show, respectively, that (i) the positions of the sample and reference standards and (ii) the method of calibration to reflectance both impact the absolute emissivity values retrieved, even on a single setup. At KCL, higher SEs are retrieved when using the substitution method of calibration (KCL_sub) than when using the comparative method of calibration (KCL_comp), with the latter the closest of all laboratory measurements to the mean of all measurements over 8–14 µm, as shown in Table 4. At CSIRO, all the SE measurements made using the Bruker VERTEX 80v FTIR spectrometer were derived using the comparative method, with sample and reference simultaneously mounted. However, merely changing the position of the sample and reference targets changed the derived SE of every samples. This can be seen by the differences between the derived SEs with the reference target in the lower port and the sample in the top port (CSIRO BG-B_S-T) and the derived SEs with the reference target as the internal wall and the sample in the top port (CSIRO BG-W, S-T) in Figures 7 and 8. CSIRO's highest SE values—and those most in agreement with the other laboratory measurements—are recorded for both artificial samples with the reference target in the lower port and the sample in the top port (CSIRO

BG-B_S-T), while the lowest emissivities for Sample 1 are found with the reference target as the internal wall and the sample in the top port (CSIRO BG-W, S-T), and for Sample 2 with the reference target as the internal wall and the sample in the bottom port (CSIRO BG-W, S-B). Due to the nature of the distilled water sample, only measurements with the water sample placed under the bottom port were possible at CSIRO with the Bruker VERTEX 80v. Of these, the SEs derived using the internal wall as reference (CSIRO BG-W, S-B) are consistently slightly higher by approximately 0.008 than those measured with the reference target in the top port (CSIRO BG-T, S-B; Figure 5).

3.2.2. Differences in Spectral Shapes between Emissivity Measurements

The spectral shapes of the SE measurements made at the different laboratories are generally consistent for the Lambertian sample (Sample 2) and for the distilled water measurements, despite absolute differences in retrieved SE. However, this is not the case for the specular sample (Sample 1), where wavelength shifts of up to 0.09 μm are evident in the 9.8–11 μm spectral region as shown in Figure 9 for a subset of the laboratories (Optosol, UT-ITC and ONERA). The wavelength outputs from FTIR spectrometers are known to have some variation resulting in wavenumber calibration procedures as in [42,43]. However, shifts of this magnitude are larger than would be expected from this. While they could suggest alignment issues within the FTIR, if that were the case we would see shifts for all samples—generally the wavelengths would be shifted by a constant correction term dependent on the laboratory setup. Given that the shifts are observed for the specular sample only and that they are observed in measurements on the same setup (Figure 10), incorrect calibration of wavelength outputs is therefore determined not to be the cause.

A more likely cause of the shifts is the different incident angles in each method. This is because, for a specular sample, the resonance wavelength will change with the incident angle, assuming a cavity effect due to the thin layer coating. For example, the ONERA measurement sequence includes a tilted beam measurement for the compensation of the sphere substitution error whereas the beam was on the normal (0°) in the Optosol setup. This theory is supported by the fact that similar shifts are observable over that wavelength range between different measurements at CSIRO that were made using the VERTEX 80v spectrometer with different sample and reference positions (Figure 10). Here the measurements made with the sample in the bottom port (which had good agreement with each other) appear to be up to 0.13 μm out of phase with those where the sample was in the top port. No spectral shifts were observed between the KCL measurements of Sample 1 with different permutations (KCL_sub and KCL_comp, Figure 7).

Other potential causes of the wavelength shifts could be changes in the water vapour and CO_2 conditions between the sample and reference measurements, non-uniformity in the PE film structure and thickness for Sample 1, different sample orientations at the time of measurement as in [44], or differences in the spectral data interval as detailed in [29] and caused by different settings in zero-filling factors for example. To evaluate the impact of sample orientation of position of the spectral features, measurements were made of Sample 1 at KCL at different orientations (0–315° in increments of 45°). The locations of the spectral features in this spectral region agreed between measurements at 0°, 90°, 180° and 270° but small wavelength shifts of approximately 0.04 μm were observed between these measurements and the measurements at 45°, 135°, 225° and 315° (which were in agreement with each other). It is likely therefore that the different sample orientations or differences in the illumination angles used within the different measurement setups could therefore at least partly explain the spectral shifts observed.

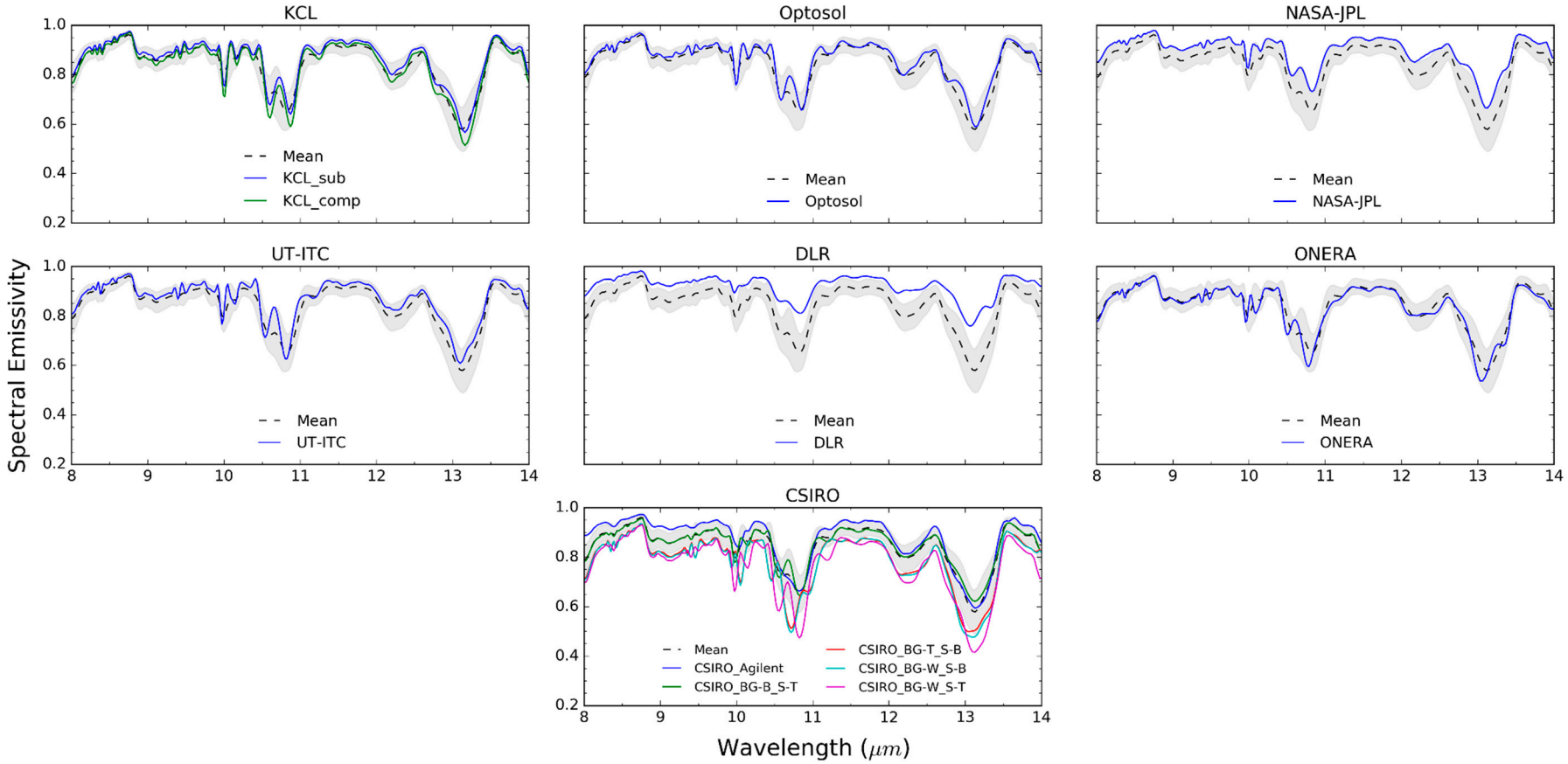

Figure 7. Measured SEs for Sample 1 from each participating laboratory, derived at a 4 cm^{-1} spectral resolution across 8–14 μm and presented against the mean and standard deviation of all SE measurements, shown as the black dashed line and grey shaded area respectively.

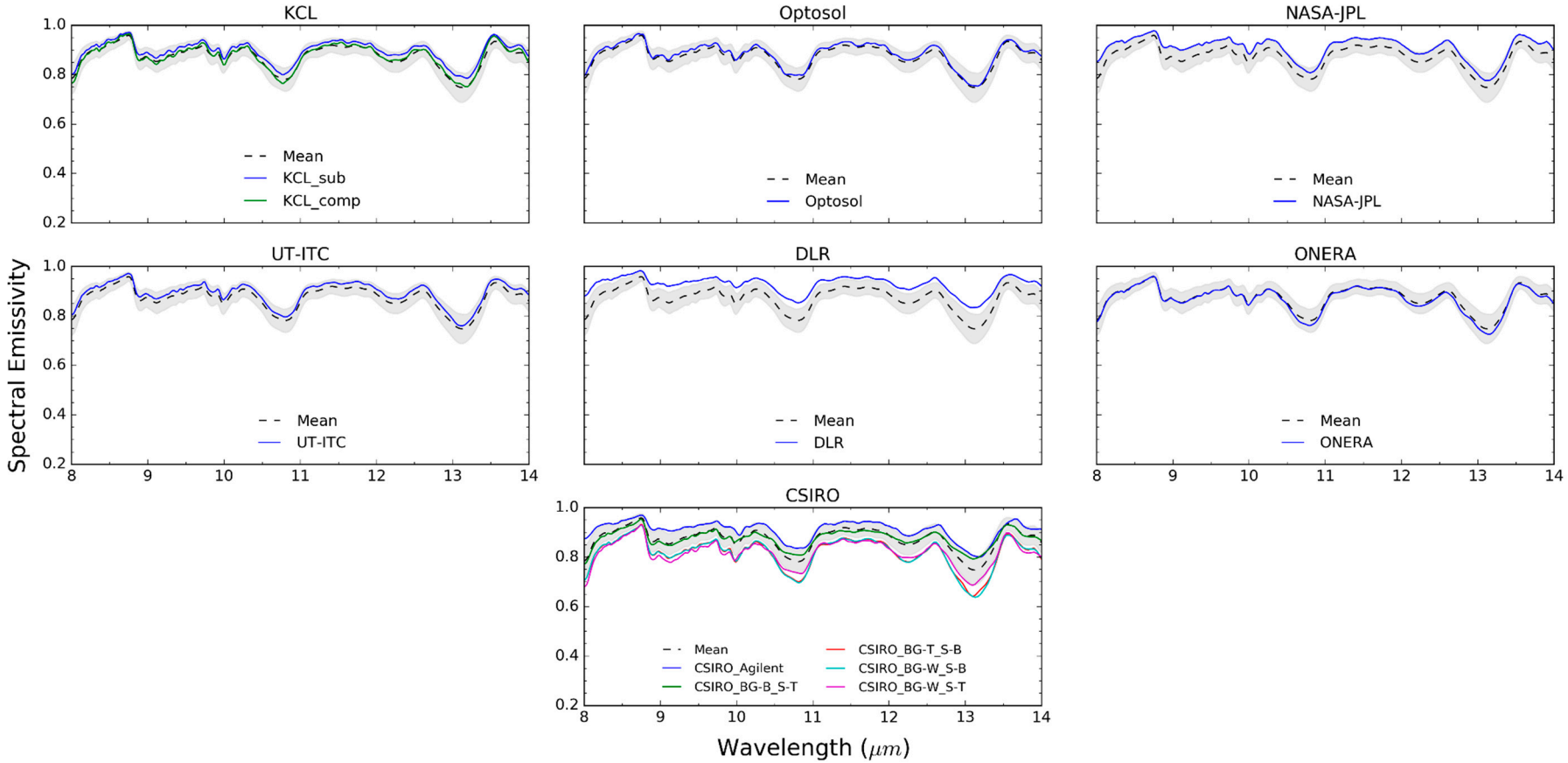

Figure 8. Measured SEs for Sample 2 from each participating laboratory, derived at a 4 cm^{-1} spectral resolution across 8–14 μm and presented against the mean and standard deviation of all SE measurements, shown as the black dashed line and grey shaded area respectively.

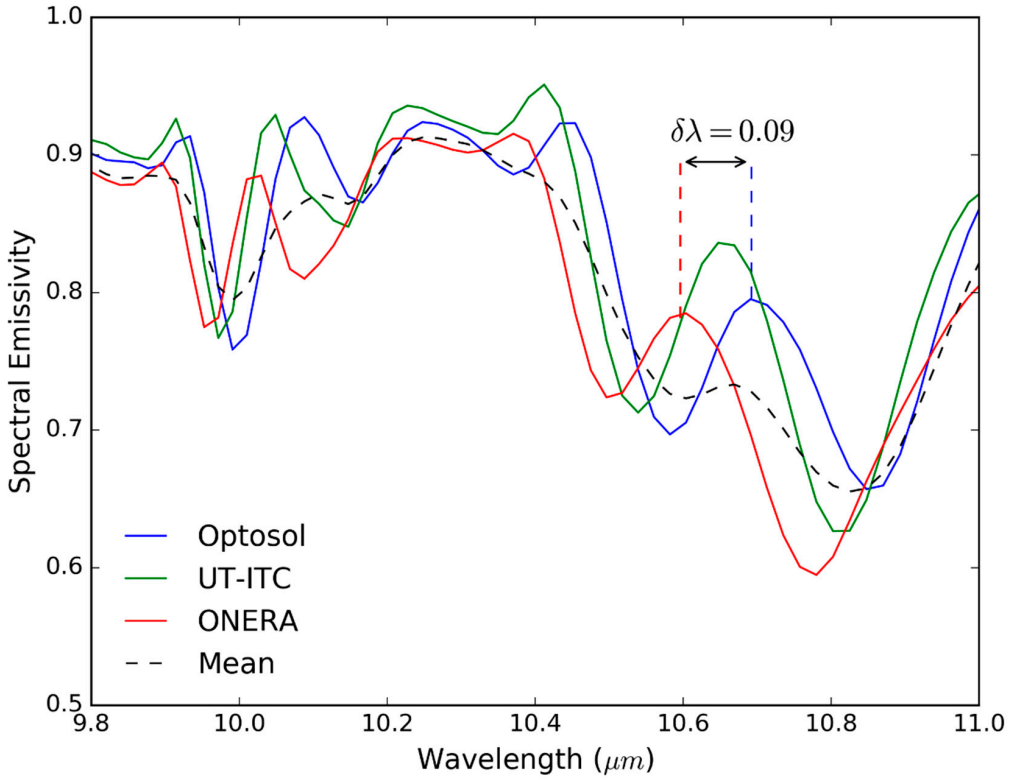

Figure 9. A subset of the laboratory SE measurements of the specular sample (Sample 1), highlighting the wavelength shifts that appear in the 9.8–11 µm spectral region. The mean SE of all measurements is also shown.

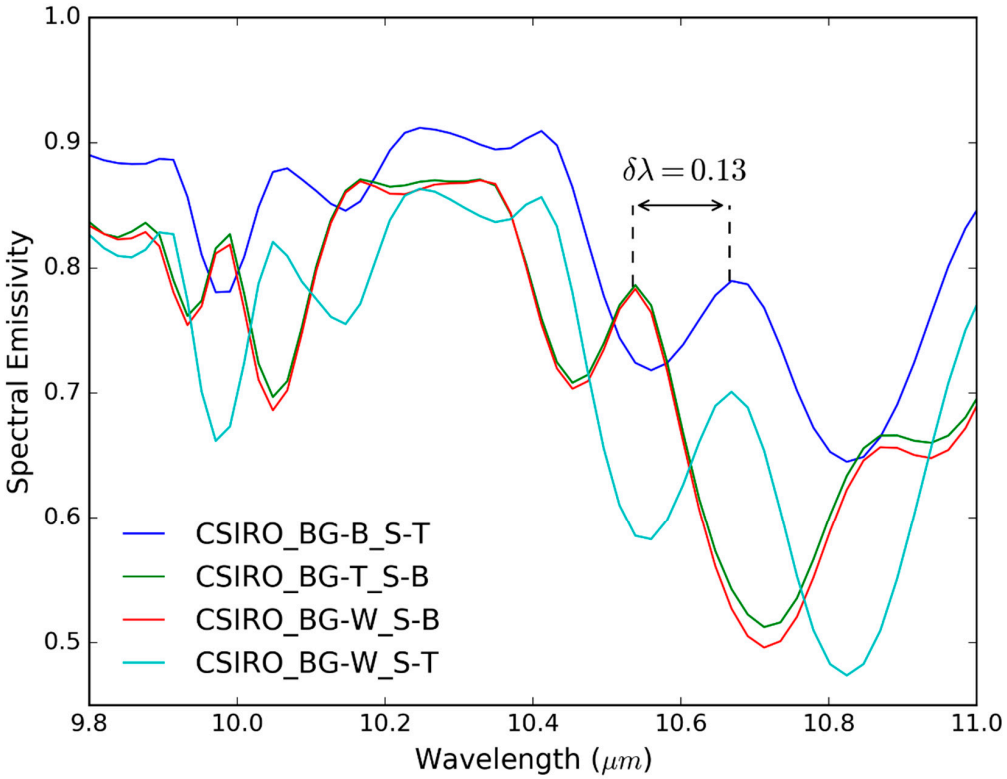

Figure 10. The CSIRO laboratory SE measurements of the specular reference sample (Sample 1) made using their Bruker VERTEX 80v FTIR spectrometer setup with the sample and reference in different positions, showing the wavelength shifts that appear in the 9.8–11 µm spectral region. For legend abbreviations see Table 2.

3.3. Comparison between Different Laboratories' Reference Standards

To identify whether the cause of the differences could be attributed to different reference standards used in the substitution approach, a comparison of SEs calculated using two different laboratories' reference standards was conducted, shown in Figure 11. Using ONERA's reference standard (with absolute reflectance provided by ONERA) within the KCL setup in the substitution mode reduces the differences between the measured emissivities of the artificial samples by between 10 and 50%. However, it does not equalise them, with the KCL measured emissivities—including those derived using the ONERA reference sample—higher than the ONERA-measured emissivities for both samples between 8 and 14 µm (with the exception of the minima around 13 µm). It also does not account for the wavelength shift between certain spectral features of Sample 1 detailed in Section 3.2.2.

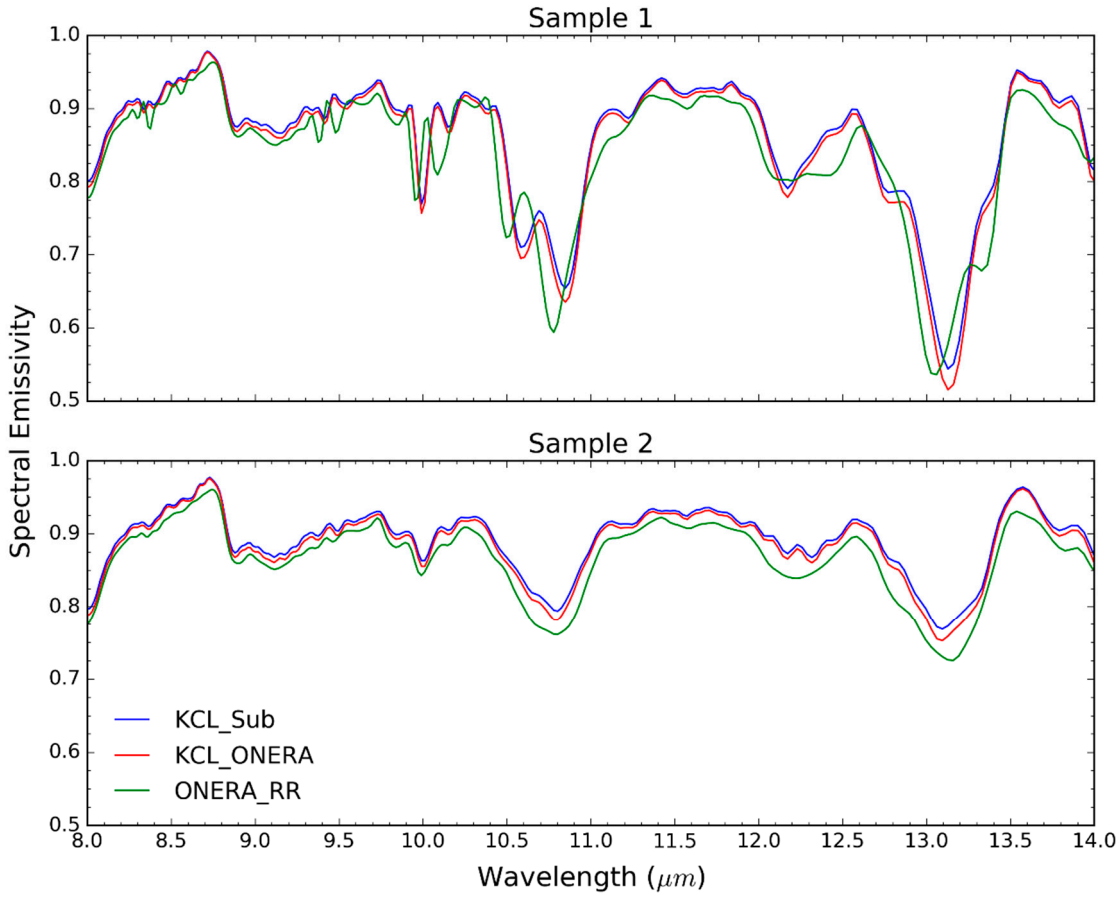

Figure 11. Surface spectral emissivities of the artificial samples Sample 1 and Sample 2, shown over the 8–14 µm spectral region and measured at KCL using the substitution method with the KCL reference standard (KCL_Sub) and the ONERA reference standard (KCL_ONERA), and at ONERA using the substitution method with the ONERA reference standard (ONERA_RR).

3.4. Implications of SE Differences for LST Retrieval

To demonstrate the impact of the SE differences we observed on typical estimates of remotely sensed LST, spectrally integrated surface emissivities (ε_i) for the artificial reference samples and distilled water were calculated across each the five thermal infrared (TIR) bands of the spaceborne ASTER instrument, which is commonly used for LST determination [45–47]. ASTER bands 10–14 are centred at 8.3 µm, 8.7 µm, 9.1 µm, 10.6 µm and

11.3 µm respectively in regions of high atmospheric transmittance as shown in Figure 12. Band integrated emissivities were calculated using:

$$\varepsilon_i = \frac{\int_{\lambda_0}^{\lambda_1} Sr_\lambda(i)\varepsilon(\lambda)d\lambda}{\int_{\lambda_0}^{\lambda_1} Sr_\lambda(i)d\lambda} \quad (3)$$

where i indicates the band number, λ wavelength, $\varepsilon(\lambda)$ the measured spectral emissivity and λ_0 and λ_1 the lower and upper bounds of the band and Sr_λ is the spectral response function for each band.

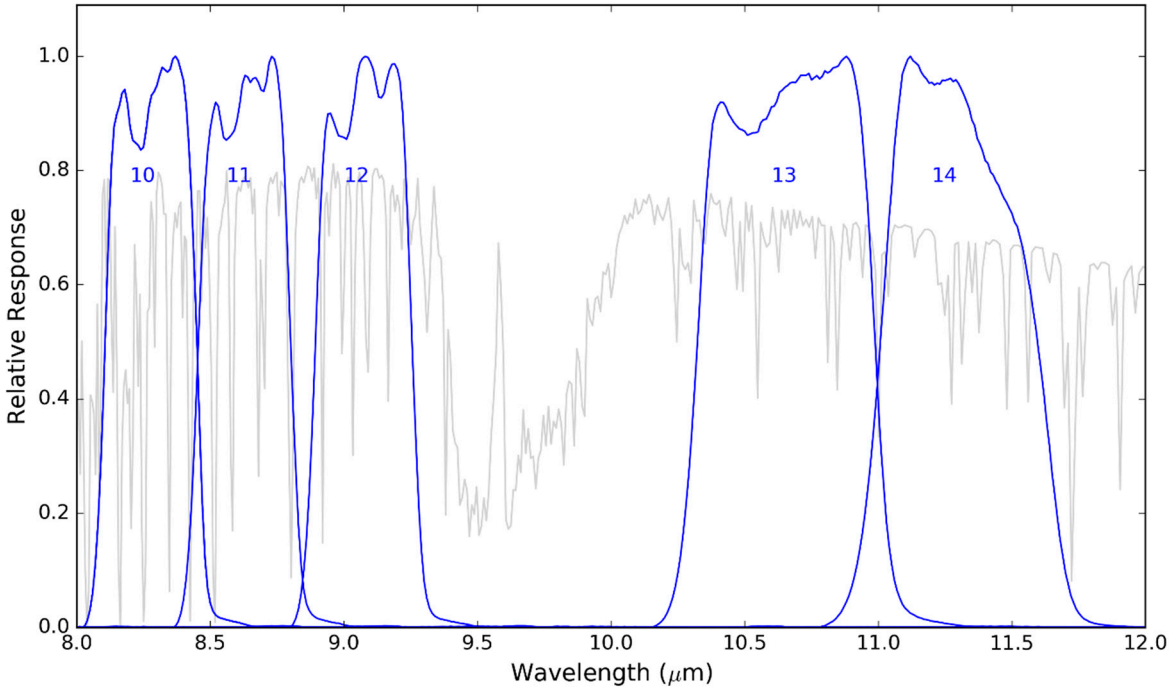

Figure 12. Spectral response functions of ASTER Bands 10–14 (blue) and the atmospheric transmittance of a mid-latitude summer atmosphere calculated using MODTRAN 5 (grey) [48].

A typical mid-latitude summer situation was simulated, assuming observation of the each of the samples with a land surface brightness temperature (BT) of 300 K, a sky BT of 260 K and a very near-surface remotely sensed observation (to ensure negligible atmospheric transmissivity and path radiance effects, and thus focus on the surface emissivity impacts only). LSTs corresponding to each ASTER band 10–14 were calculated as in [8], such that the LST of band i was equal to:

$$LST_{\mathrm{i}} = B_{\mathrm{i}}^{-1}\left[\frac{1}{\varepsilon_{\mathrm{i}}}\left(L_{\mathrm{surf,\,i}} - (1-\varepsilon_{\mathrm{i}})L^{\downarrow}_{\mathrm{sky,\,i}}\right)\right] \quad (4)$$

where ε_i is the surface emissivity in band i coming from Equation (3), $B_{\mathrm{i}}^{-1}(L)$ is the inverse Planck function describing the blackbody equivalent temperature T (kelvin) of spectral radiance L_{i} ($\mathrm{W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}}$) in band i, $L_{\mathrm{surf,\,i}}$ the spectral radiance ($\mathrm{W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}}$) corresponding to the surface BT in band i and $L^{\downarrow}_{\mathrm{sky,\,i}}$ the spectral radiance ($\mathrm{W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}}$) corresponding to the sky BT in band i.

Figure 13 shows the statistical distribution of LSTs calculated using Equation (4) by sample and by ASTER TIR band, with the box showing the interquartile range and whiskers the distribution (excepting outliers). LSTs derived using the convolved artificial sample emissivities range by over 3.5 K in each band, with a maximum difference of 17.8 K (Sample 1, Band 13). These differences greatly exceed both the GCOS target accuracy and currently

achievable requirements for LST as an ECV, which are 1 K and 2–3 K respectively [49]. The difference of nearly 20 K for Sample 1 in Band 13 is related to this waveband covering the spectral range containing the observed wavelength shifts in the Sample 1 SE measurements, and being an area of increased atmospheric attenuation and thus stronger downwelling irradiance impact.

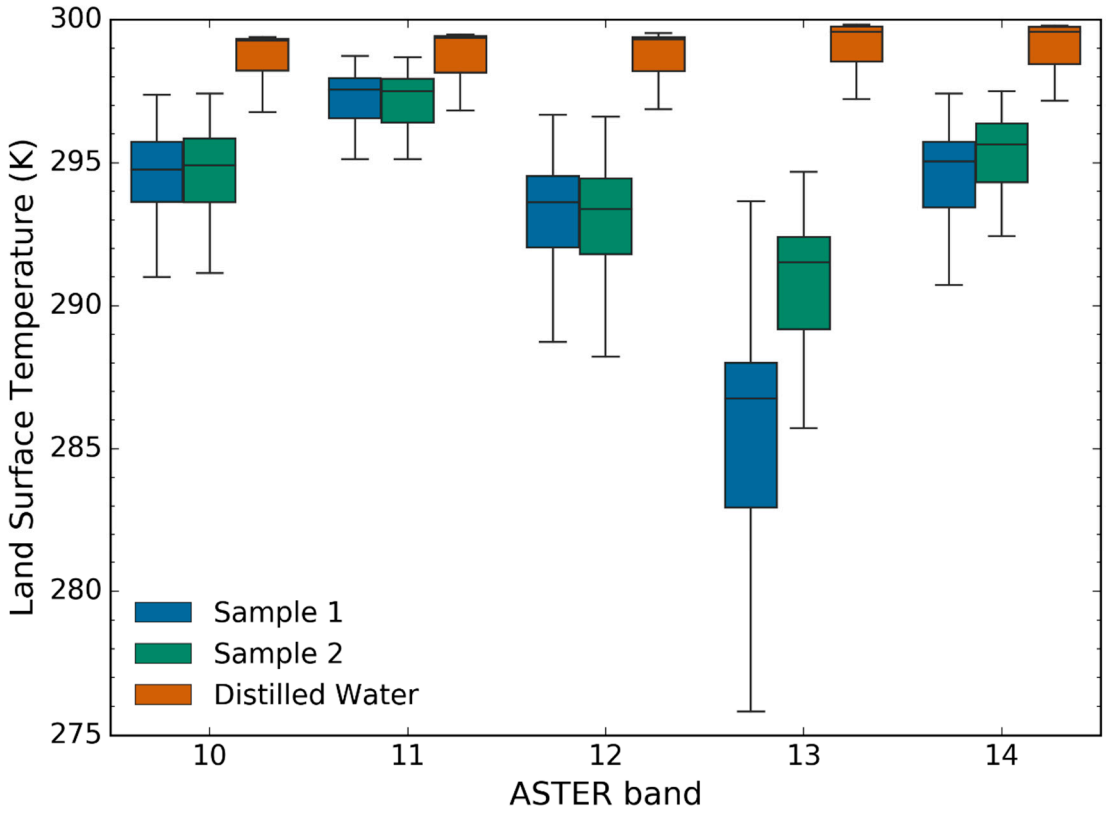

Figure 13. Range of land surface temperatures (K) retrieved from simulated brightness temperature data calculated assuming a typical mid-latitude summer scene with a surface brightness temperature of 300 K, a sky temperature of 260 K and the measured emissivities for each of the three reference samples convolved to ASTER TIR bands 10–14.

The range of the LSTs calculated using the distilled water emissivities is reduced compared to those calculated using the artificial sample emissivities, reflecting the greater agreement between the SE measurements of distilled water. However, the range of the LSTs calculated using distilled water spectra convolved to the ASTER TIR bands is still 2.5 K in all bands, thus still exceeding the GCOS target accuracy requirements.

4. Discussion

Absolute differences observed between the different measurements were larger than anticipated with no clear cause. The derived emissivity does not correspond to different spectrometer types as may be expected: DLR and CSIRO's measurements were both made on setups based on a Bruker VERTEX 80v FTIR spectrometer, albeit with different spheres, however there were large differences between these measurements for all three samples over the full wavelength range. Furthermore, the measured absolute differences cannot be solely attributed to the use of the reference standard, although results from Section 3.3 indicate that the uncertainty in reference standard calibration is a key factor in the SE uncertainty. The contribution to uncertainty from reference calibration is particularly pertinent given that Labsphere Infragold standards can no longer be bought with NIST traceability calibration certificates. Differences in the absolute reflectance of the reference standards could be due to different coatings but could also result from physical damage or degradation from humidity absorption. To reduce uncertainty regarding the latter,

laboratories should ensure regular calibration of their reference standards to monitor for drift.

Results from all three samples considered indicated that three of the four CSIRO measurements made using the VERTEX 80 had a consistent negative bias, with the measurement made with the reference target in the lower port and the sample in the top port (CSIRO BG-B_S-T) the only one in agreement with the others. Given the similarity in the spectral shapes measured by this setup and those measured at other laboratories for all three samples, this could suggest that CSIRO may need to re-characterise their reflectance standards for the other three configurations used with the VERTEX 80v spectrometer. However, these differences may also be due to the different optical path lengths of the sample beam with each permutation, or to directional reflectance effects of the sample at the different incident angles of each permutation. Further investigation is recommended if CSIRO wish to use any of the three biased configurations (e.g., to measure liquid samples, which is not currently possible with the sample in top port configuration). Removing these three measurements from the analysis reduces standard deviations to $\pm$0.089 (14.69%), $\pm$0.038 (5.16%) and $\pm$0.008 (<1%) for Sample 1, Sample 2 and distilled water respectively across the 2.5–14 µm wavelength range. Furthermore, the impacts on LST are reduced considerably without these measurements, with the range of the LSTs calculated using distilled water spectra convolved to the ASTER TIR bands reduced to <0.45 K in all bands.

Differences were also observed between measurements made on the KCL setup with the reference target in different positions. Possible causes for these differences at KCL are (i) differences in the path lengths for each measurement setup that remain unaccounted for as discussed in Hecker et al. [27], (ii) insufficient correction for the substitution error in the substitution method, (iii) differences in reflective properties of the reflectance targets (being flat and curved respectively for the substitution and comparative methods) and (iv) incorrect characterization of the reference target spectral reflectance (the Infragold reference panel and sphere wall for the substitution and comparative methods respectively). In terms of the latter possible cause, it was identified that scaling the provided absolute spectral reflectance of the internal wall of the KCL integrating sphere (which is used as the reference target during the comparative method) by a factor of 0.87 (so that $\rho_r = 0.84$) brought the derived SE very close to that derived using the substitution method (which used the Infragold reference panel as the reference) for all three samples, as shown over the LWIR region in Figure 14. However, the reflectance of the internal wall coating should be much higher than this given the material type and the fact that both this system and the Infragold target used as the reference target for the substitution method are new. Note that it is not physically possible for the quoted spectral reflectance of the Infragold reference panel to be higher by this amount as this would result in emissivities above 1. Evaluating the performance of the two methods provides mixed results. While measurements derived using the comparative method of calibration were found to be closer to the mean for all samples, KCL's distilled water emissivity measurements retrieved using the substitution method of calibration were observed to be closer to the ECOSTRESS spectral library spectrum than those derived using the comparative method. This could indicate that the mean may have been negatively skewed by the afore-mentioned bias in the CSIRO measurements. Discounting the three negatively biased measurements from CSIRO from the analysis supports this: KCL's measurements derived using the substitution method were in good agreement (<0.01) with the recalculated mean over 2.5–14 µm for all three samples while KCL's measurements derived using the comparative method were in poorer agreement, with differences of up to 0.04 across 2.5–14 µm. Based on these results, further investigation should be conducted to determine the cause of the differences between the measurements from KCL and to determine an optimal approach.

The increased variability in the MWIR than LWIR observed is likely due to the increased atmospheric effects in this region, with the DLR measurements of Samples 1 and 2 clearly impacted in the CO_2 region (Figure 4). An alternate explanation for the reduced variability in the LWIR for both artificial samples could be because this is an area of high

emissivity (and thus low reflectance), which Hecker et al. [27] observed to be areas of better agreement in their intercomparison of emissivity spectra from different laboratories.

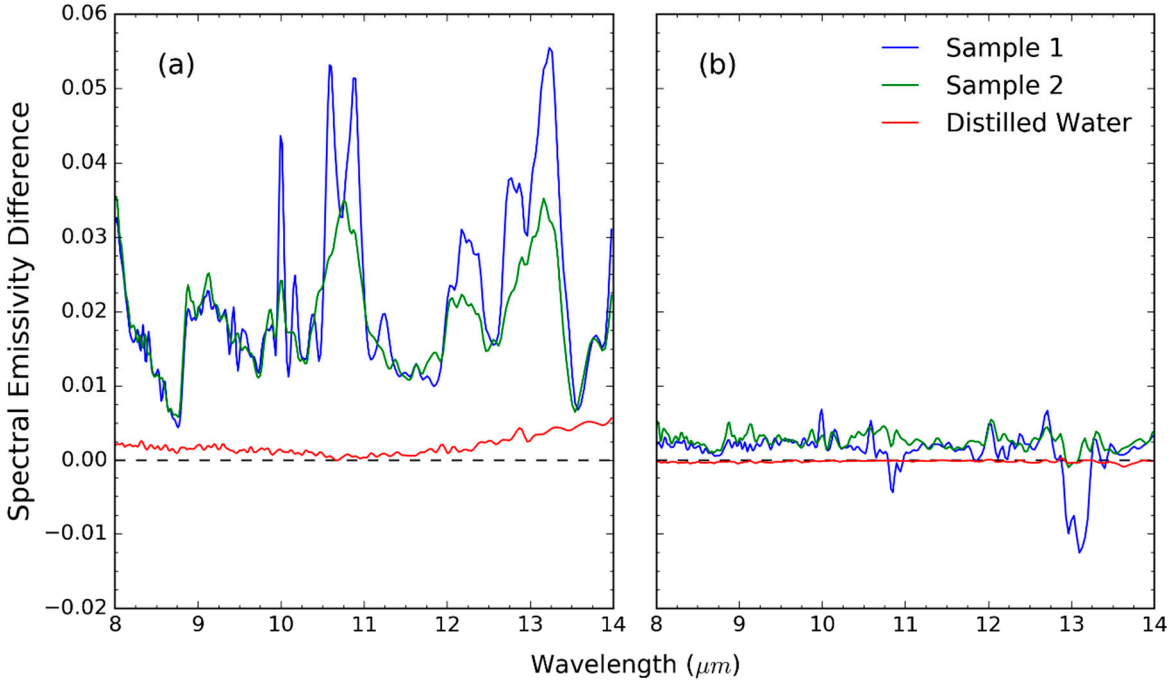

Figure 14. Differences between the surface spectral emissivities derived at KCL using the substitution method and the comparative method ($\varepsilon_{sub} - \varepsilon_{comp}$), shown here over 8–14 µm for both the artificial samples and distilled water when (**a**) leaving the reflectance of the comparative reference target as specified by the manufacturer and (**b**) scaling the absolute reflectance of the comparative reference target by 0.87.

This latter interpretation could also be why the distilled water measurements (with uniformly high emissivity) had reduced variability compared to the artificial sample SE measurements (which had variable emissivities between 0 and 1). However, it is more likely that the increased variability of the SE measurements of the artificial samples was due to the composition of these samples and their interaction with different setups. Tsilingiris [50] provide a transmittance spectrum for polyethylene (PE), and considering this against the measured spectra of Samples 1 and 2 in Figure 4, it is clear that variability amongst the different measurements was lower in the regions where PE had a low transmittance (3.5 µm, 6.9 µm and 13.8 µm). In all other spectral regions, the PE formed a multilayer system, which is potentially sensitive to directional illumination characteristics. Differences in the incident angles upon the samples within the different measurement setups could therefore at least partly explain the SE variations seen. Given that Sobrino and Cuenca [51] observed that emissivities in field measurements tended to decrease with increases in the observation angle, results from this study indicate that future work should be conducted to explore whether emissivities from DHR setups in the laboratory similarly correlate with the incident angle. Materials with expected directional behaviour should in particular be considered.

The observed spectral shifts between different measurements of the specular sample also raise interesting questions about the impact of incident angles on SE measurements. While this may not be an issue for non-specular samples without coating (and therefore for most natural samples), these discrepancies indicate that further work should be conducted to confirm and investigate the impact of the incident angle on the spectral stability and absolute emissivity. Spectral shifts of the magnitude observed in this study will have more implications when working with data from hyperspectral rather than multispectral thermal imagers. Conversely airborne hyperspectral instruments such as NASA-JPL's airborne

Hyperspectral Thermal Emission Spectrometer (HyTES) sensor have so many narrowband (18 nm) spectral channels between 7.5 and 12 μm that TES and spectral smoothness approaches can be applied, which reduce the need to prescribe emissivity in advance [52,53]. However, such TES approaches often rely on laboratory emissivity spectra to derive empirical relationships used within the algorithm, and so their accuracy is still important even though for each hyperspectral image pixel the emissivity is directly retrieved [10]. Such wavelength shifts may also have implications for use of spectral emissivity features in, e.g., mineral identification studies [54], and also may affect in situ LST measurements given that radiometers commonly used in LST validation studies are affected—such as the Heitronics KT15.85 IIP radiometer with a spectral range of 9.6–11.5 μm [16,55].

Consideration of how the uncertainty of individual SE measurements translates into retrieved LSTs in Section 3.4 indicates the importance of reducing uncertainties in laboratory SE measurements to improve remotely sensed LST estimates. While it must be acknowledged that the artificial samples are not representative of many land surfaces, with most natural surface less reflective in the LWIR, similar samples may be observed in remote sensing of urban areas (e.g., for urban heat island monitoring) or monitoring of plastic pollution in water [7,56,57].

5. Summary and Conclusions

Surface spectral emissivity data collected with twelve different laboratory spectrometer measurement setups made at seven different laboratories were compared over the MWIR and LWIR spectral ranges in a Round-Robin intercomparison exercise. All measurements were based on the principle of illuminating the sample with an intense source of infrared radiation, measuring the reflected signal, and converting this to an emissivity spectrum using Kirchhoff's Law. Three different samples were used for the exercise. The first two were artificial samples constructed from gold and aluminium sheets each laminated in PE films that had Lambertian and specular characteristics respectively and with widely varying emissivity features across the 2.5–14 μm spectral range. The third sample was distilled water, which has a relatively flat emissivity spectrum close to unity.

Comparing the measurements from the different laboratories we found that the inter-setup variability of the SE measurements was larger than anticipated, with differences in magnitude and spectral shape. Standard deviations of ±0.092 (15.98%) and ±0.054 (7.56%) were identified across the 2.5–14 μm spectral range for Samples 1 and 2 respectively. Repeated measurements using the same measurement setup and protocol at different times confirmed that observed SE differences were not attributable to changes in the sample properties over the course of the study but were rather due to the different setups and measurement procedures used in the various laboratories. Variability was greater in the MWIR rather than LWIR spectral region, likely due to differing efficiencies of atmospheric purging, which impact this region more. SE differences across the LWIR atmospheric window (8–14 μm), which is the most important for the remote sensing of LST, were ± 0.046 and ± 0.037 respectively for Samples 1 and 2, and most of the Sample 2 SE measurements were within 0.02 of the mean. The greater variability for the specular sample (Sample 1) over this region was attributed to spectral shifts, with differences between identification of spectral maxima and minima of up to 0.09 μm between different laboratories and up to 0.13 μm between different positional permutations on one setup were observed in the 9.8–11 μm spectral range in particular. Investigation indicated potential causes of these spectral shifts to be different sample orientations during measurements or differences in the incident angles within the different measurement setups. The latter cause was also identified as a potential cause of the absolute emissivity differences. Further investigation is therefore recommended into the impact of directional effects in laboratory measurements of emissivity (particularly for materials with known directional behaviour) given recent advances into understanding the angular dependence of emissivity for field and satellite measurements [6,58].

Use of different reference standards was found to contribute to the observed SE differences between different laboratory measurements but not to be the sole factor. Nonetheless, the differences observed from use of a different reference standard suggest that uncertainty in the reference standard calibration is a key factor in emissivity uncertainty in laboratory measurements. Regular calibration of the reference standards is recommended to reduce this uncertainty.

SE variability was comparatively lower for distilled water than for the artificial samples, with a mean emissivity of 0.962 $\pm$ 0.028 determined over the 2.5–14 μm spectral range. These uncertainties are larger than those observed in other studies but standard deviations were reduced to 0.008 when discounting measurements from one laboratory with a consistent negative bias (likely indicating inaccurate calibration of their reference standard). Other contributions to the SE differences included the method of calibration to reflectance, different incident angles and the placement of the sample and reference standards within the same setup. Regarding the setup at KCL, the primary laboratory used in this study, the measurements indicate that the comparative method of calibration was closer to the mean of all measurements. However consideration of the distilled water suggested instead improved performance using the substitution method and therefore further investigation is recommended.

The impact of the determined spectral emissivity differences on LST retrieval was evaluated by considering a typical mid-latitude summer scene with the surface emissivities set to the sample emissivities convolved to the ASTER TIR bands. With Sample 2 (the diffuse sample) considered more representative of natural surfaces than Sample 1, use of the three samples in this simulation provided LST error estimates over diffuse surfaces, over an extreme case of 100% specular surfaces and over water bodies. Calculated LSTs using the convolved artificial sample emissivities ranged by over 3.5 K in each band, with a maximum difference of 17.8 K (Sample 1, Band 13). The range of the LSTs calculated using distilled water spectra convolved to the ASTER TIR bands (and thus an error estimate for retrieval of LST over water bodies) was lower, but still around 2.5 K. The variability of the artificial samples and the distilled water emissivities measured at different laboratories would therefore result in uncertainties in LST estimates that exceeded the target accuracy requirements for satellite observations of land surface temperature, even without considering contributions from atmospheric effects.

Overall, our findings highlight the need for the infrared spectroscopic community to work towards standardized and interlaboratory comparable results, with regular calibration of reflectance standards and the laboratory setup against SI traceable standards.

Author Contributions: Conceptualization, M.F.L., M.W., J.J.H., M.K., C.H. and S.J.H.; Data curation, M.F.L., J.J.H., M.K., C.H., E.A., W.R.J., A.M., L.P. and I.C.L.; Formal analysis, M.F.L.; Investigation, M.F.L., J.J.H., M.K., C.H., E.A., W.R.J., A.M., L.P. and I.C.L.; Methodology, M.F.L., M.W. and L.P.; Project administration, M.F.L. and M.W.; Resources, M.K. and F.B.; Supervision, M.F.L.; Visualization, M.F.L.; Writing—original draft, M.F.L. and M.W.; Writing—review and editing, M.F.L., M.W., J.J.H., C.H., S.J.H., L.P. and I.C.L. All authors have read and agreed to the published version of the manuscript.

Funding: Support for this research came partly from NERC National Capability funding to the National Centre for Earth Observation (NE/Ro16518/1).

Acknowledgments: We thank Chris Stapleton (Bruker), Francis O'Shea (King's College London) and Benjamin Gridley (Queen Mary University of London) for their support in setting up the laboratory spectrometer setup at King's College London.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. National Centre for Earth Observation at King's College London (KCL)

The setup deployed at NCEO (National Centre for Earth Observation) King's College London (KCL) was based on that at the University of Twente Faculty ITC (UT-ITC), described by [27]. It consists of a Bruker VERTEX 70 FTIR spectrometer modified with an

external liquid-nitrogen-cooled HgCdTe (mercury-cadmium-telluride (MCT)) detector and an external integrating sphere of diameter 150 mm with an interior diffusely reflecting gold coating and a 30 mm sampling port to enable directional-hemispherical reflectance measurements of large, inhomogeneous samples (Figure A1). There are two internal sources (an air-cooled globar for the MIR and a tungsten lamp for the NIR) and an external high-power water-cooled globar for the MIR. The external infrared source was used for this study for improved signal-to-noise ratio. In addition to the MCT detector, the sphere is equipped with an external InGaAs detector for consideration down to the NIR, thus enabling spectral measurements from 0.7 to 16 µm. The entire system (including the integrating sphere) is continuously purged with H_2O- and CO_2 free air at a flow rate of at least 200 L/h to reduce atmospheric features in the spectra and prevent degradation of the KBr beamsplitter. The incident beam has to be convergent at an angle of 3–4°. The incidence angle is 12° in order to prevent the specularly reflected part from escaping through the entrance hole. The measured sample area is about 25 mm in diameter. Spectrometer settings used in this study are given in Table A1.

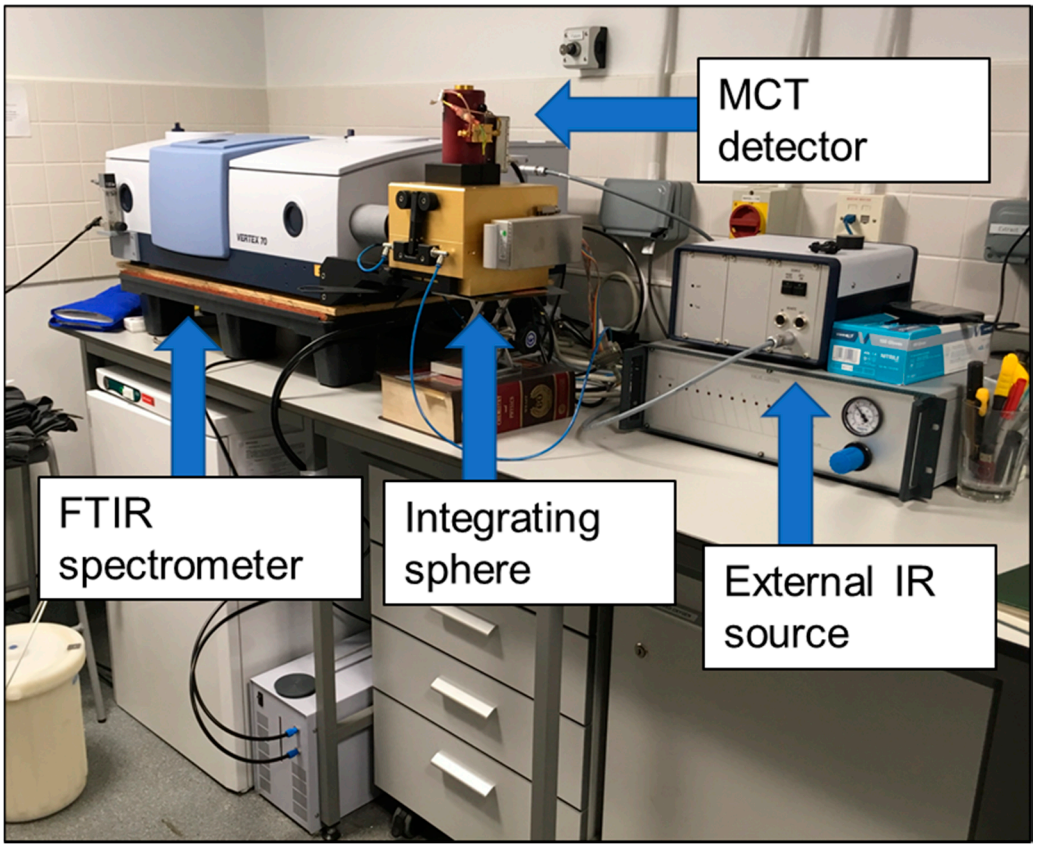

Figure A1. The VERTEX 70 spectrometer setup at King's College London with details highlighted.

Table A1. Standard measurement settings for the VERTEX 70 Spectrometer setup at King's College London.

Parameter	Settings
Spectral resolution	0.5 cm^{-1}, 4 cm^{-1} and 8 cm^{-1}
Spectral recording range	2.5–16 µm (4000–625 cm^{-1})
Aperture setting	4 mm
Phase resolution	32 cm^{-1}
Phase correction mode	Mertz
Apodization function	Blackmann-Harris 3-Term
Zero-filling factor	2

The design of the sphere at King's allows for both the substitution and the comparative calibration procedures, with an internal rotating mirror to enable the comparative method, as shown in Figure A2. To perform a reference measurement in the comparative mode, the folding mirror of the sphere (which also acts as a baffle) was rotated such that the incoming energy was reflected onto the gold-coated sphere wall instead of the sample. After the reference measurement, the folding mirror was rotated back and the sample in the sample port measured.

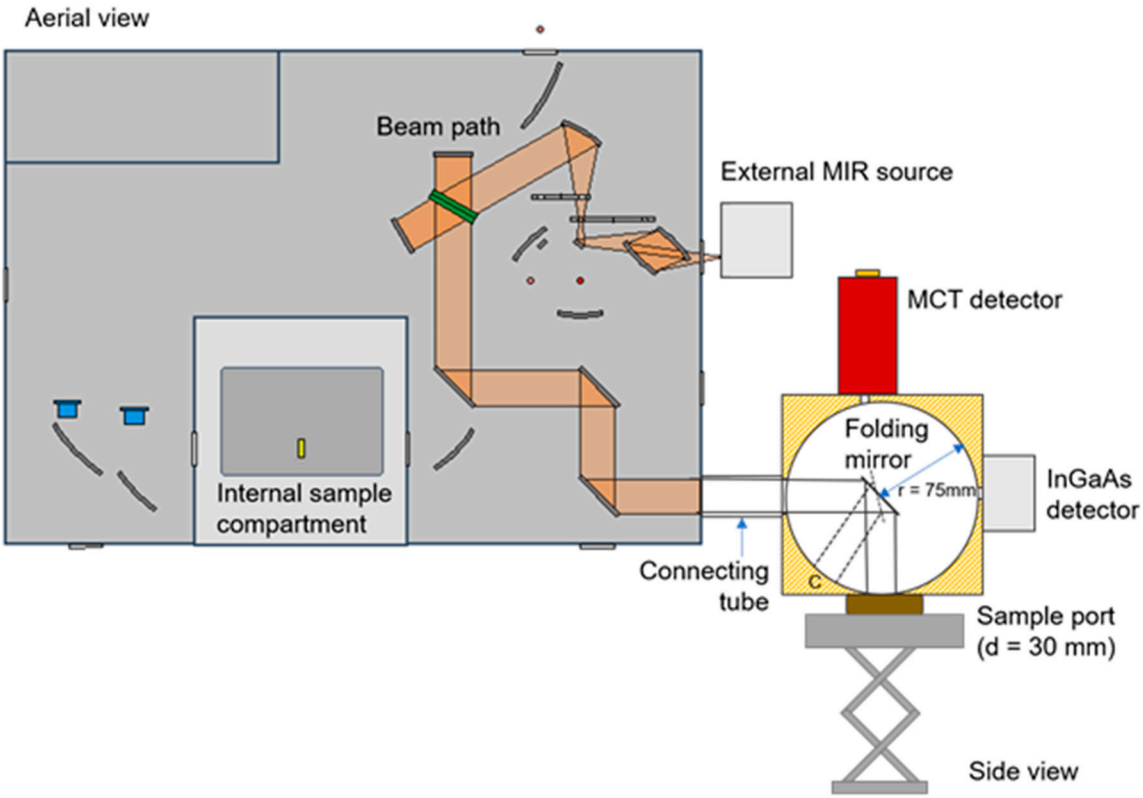

Figure A2. A sketch of the instrument setup at King's as it was used for this study, with the external MIR source deployed. The dashed beam path inside the sphere shows the beam position when the comparative calibration method is deployed, and the sphere wall is used as the reference material (marked "C" in the sketch).

Both the substitution (KCL_sub) and the comparative (KCL_comp) methods were utilised for this study, in order to analyse the performance of each method. For the substitution method, a Labsphere Infragold target (13 cm × 13 cm) with a stated reflectance of 98% above 1000 nm was used as the reference target (Figure A3b) while the internal wall of the integrating sphere (Figure A3a), with a stated reflectance of 95% above 1000 nm was the reference material for the comparative method. Hecker et al. [27] experienced problems with the comparative approach of calibration, since the slightly different path lengths provided by the two mirror positions led to strong residual influences of atmospheric gases on the resulting reflectance spectra (despite continual purging). The King's system was designed to produce identical average path lengths for the calibration and sample measurements in order to avoid this issue.

Sample reflectances (ρ_{sample}) were calculated from the single channel average scans using Equation (A1), taken from Hecker et al. [27]:

$$\rho_{\text{sample}}(\lambda) = \frac{V_{\text{sample}}(\lambda) - V_{\text{open}}(\lambda)}{V_{\text{ref}}(\lambda) - V_{\text{open}}(\lambda)} \rho_{\text{ref}}(\lambda) \quad \text{(A1)}$$

where V_{sample}, V_{ref} and V_{open} are the single channel scans in arbitrary units, and ρ_{ref} is the reflectance spectrum of the standard material (all wavelength λ dependent).

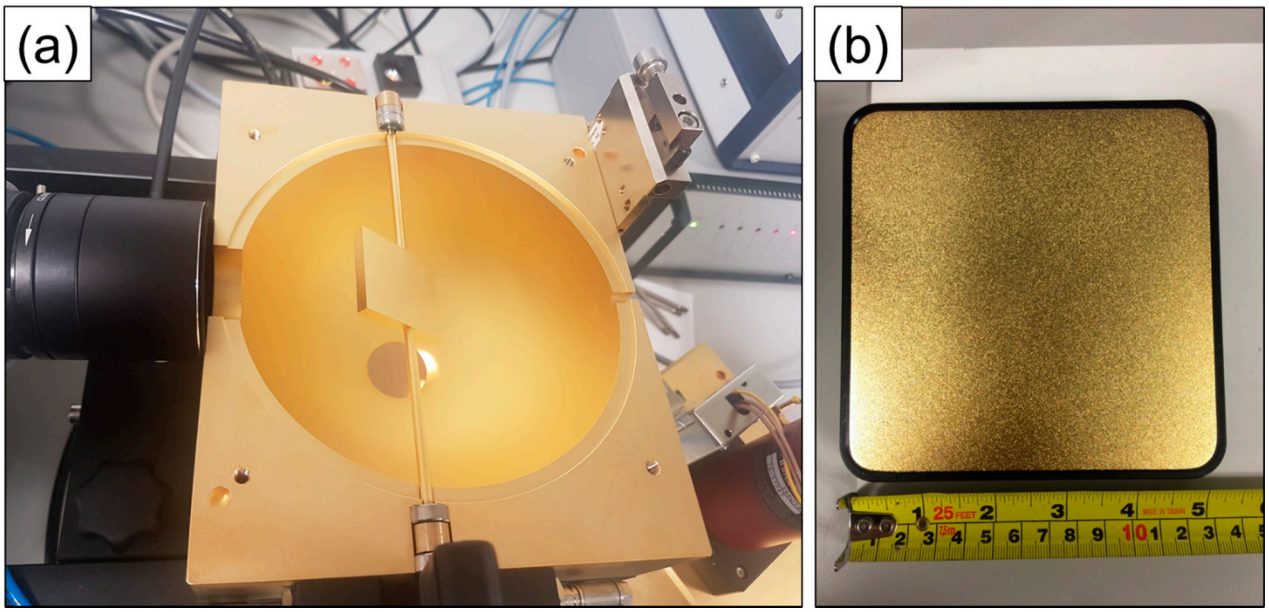

Figure A3. (**a**) The inside of the integrating sphere (image taken during installation) and (**b**) the diffuse Infragold target used as the reference target for the substitution method of calibration for the setup at King's College London.

The number of scans collected depends on the signal available for each material, with a larger number of measurements made for low signal spectra (e.g., open port or distilled water). In this study, 500 scans were made for a reference measurement, 1000 scans for an open port measurement, 500 scans for each of the artificial samples and 1000 scans for the distilled water. A full measurement sequence (sample, reference and open port) generally takes 30 minutes. Once collected, the single channel scans are coadded and averaged to provide a mean single channel measurement spectrum for each sample, reference standard and open port. A correction was applied for the substitution error as detailed in [33] as in the NASA-JPL methodology (Appendix C). Emissivity spectra (ε_{sample}) as a function of wavelength were calculated from the reflectance spectra using Kirchhoff's Law as in Equation (A2):

$$\varepsilon_{sample}(\lambda) = 1 - \rho_{sample}(\lambda) \quad \text{(A2)}$$

Appendix B. Optosol GmbH (Optosol)

The set-up applied by Optosol GmbH was based on a Bruker VERTEX 80 (Figure A4) and belonged to Fraunhofer Institute for Solar Energy Systems (ISE) in Freiburg, Germany. The samples and references were mounted simultaneously (comparative method) vertically at the infragold sphere wall (200 mm inner diameter), in contrast to the others (Figure A5). Therefore, we prepared solid state samples for the round robin. The measured area was below 20 mm in diameter. The down-looking MCT detector limited the spectral range to 1.7–17 µm. The parameters for the measurements are shown in Table A2. The specular reference and the diffuse reference were calibrated against a primary standard by NPL, Teddington, UK. Both sample and reference measurement consisted of 512 coadded scans, with the full measurement sequences typically taking 30 min.

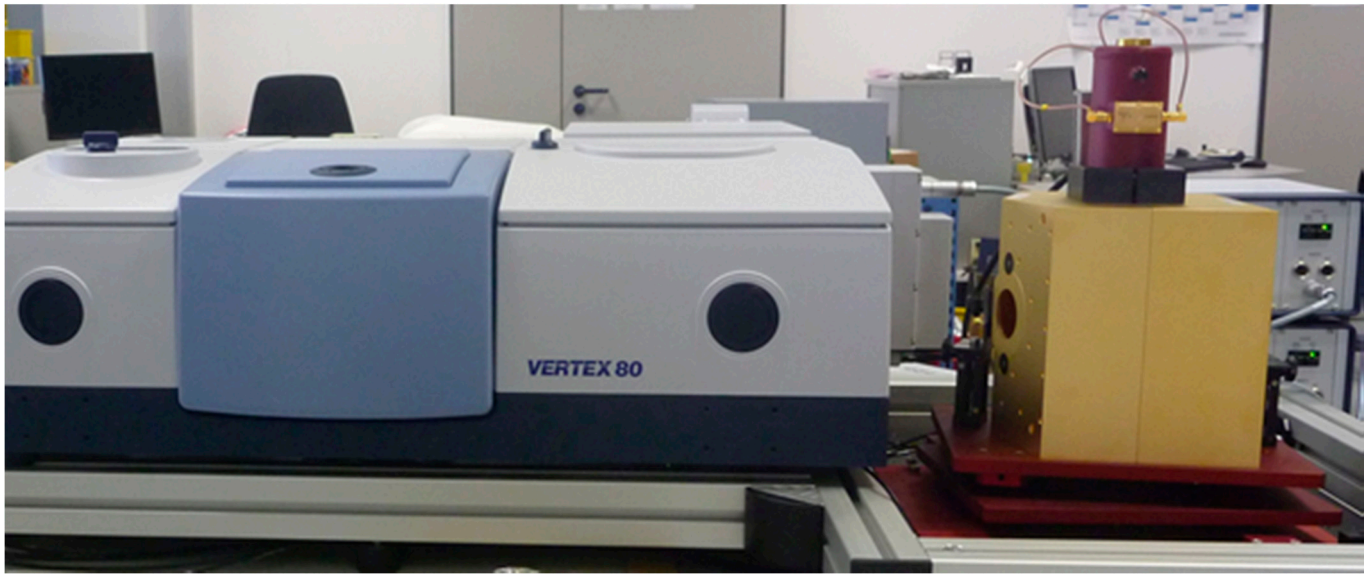

Figure A4. The set-up applied by Optosol GmbH, based on a Bruker VERTEX 80.

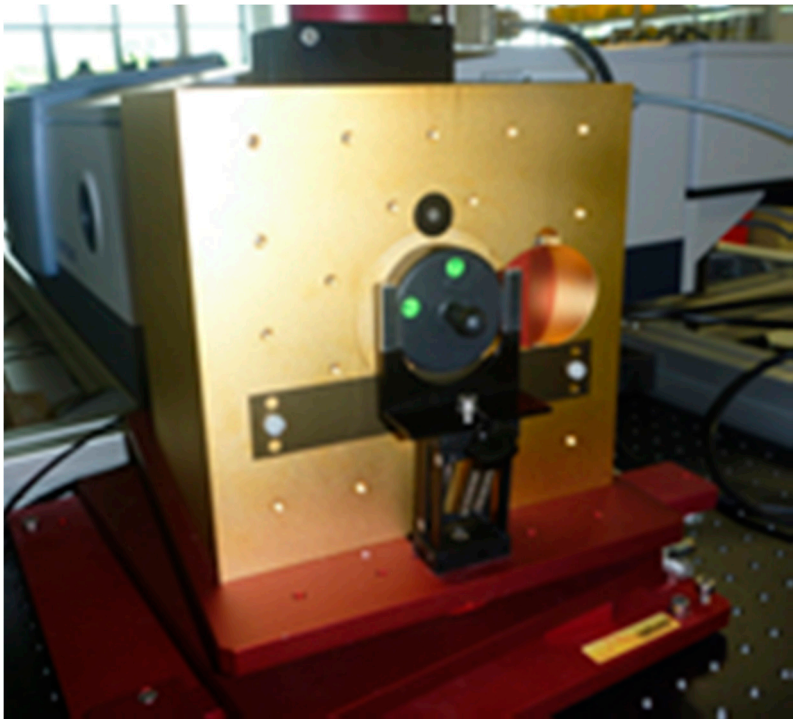

Figure A5. The integrating sphere setup at Optosol GmbH.

Table A2. Standard measurement settings for the setup applied by Optosol.

Parameter	Setting
Spectral resolution	8 cm^{-1}
Spectral recording range	1.47–17.5 µm (6800–570 cm^{-1})
Aperture setting	4 mm
Phase resolution	32 cm^{-1}
Phase Correction mode	Mertz
Apodization function	Blackman-Harris 3-Term
Zero-filling factor	16

Appendix C. NASA—Jet Propulsion Laboratory (NASA-JPL)

Measurements at the NASA Jet Propulsion Laboratory (NASA-JPL) were made with a Nicolet 6700 FTIR spectrometer equipped with a down-looking 5-inch (127 mm) diffuse gold coated integrating sphere with a 1.5 inch (38 mm) sample port manufactured by Labsphere. The reference plate is also diffuse gold (Infragold), with reflectance calculated using the substitution mode of calibration. The measurements of the two samples (1 and 2) and distilled water were made at three spectral resolutions, 0.5, 4 and 8 cm^{-1}. The usual mode of measurement is 4 cm^{-1} resolution, with a correction made for loss at the sphere at this resolution [33]. Standard measurement settings are shown in Table A3.

To avoid damage to the samples, the artificial samples were not placed directly against the sphere (distance from port <1 mm) during measurement. The instrument was purged continuously (30 psi at a flow rate of 30 cubic feet/hr). A single measurement of 512 scans was collected of the reference, while three repeat measurements of 512 scans were collected, coadded and averaged for each sample measurement to improve signal-to-noise. A full measurement sequence took around 25 min.

Table A3. Standard measurement settings for the setup at NASA-JPL.

Parameter	Setting
Spectral resolution	4 cm^{-1}
Spectral recording range	1.47–17.5 μm (6800–570 cm^{-1})
Aperture	8 mm
Phase Correction mode	Mertz
Apodization function	Happ-Genzel
Zero-filling factor	0

Appendix D. University of Twente Faculty of Geo-Information Science and Earth Observation (UT-ITC)

The instrument setup and general measurement procedures at UT-ITC are explained in depth in [27]. Differences to the published procedures were (1) N2 purge gas of 100 L/h was replaced with laboratory purge gas generator of 200 L/h and (2) the use of the substitution method for calibration with an Infragold standard, shown in Figure A6.

Twenty-four hours before that start of measurements the purge was started at 200 L/h and the external globar source was started, with both remaining on until the end of campaign. Each measurement day the MCT detector was cooled with LN_2 in the morning. After this warm start, a 1 h stabilization time was used before start of measurements. The (already cold) dewar was refilled over lunch, with a waiting time of 15 min following this.

Samples 1 and 2 and dark current (open sample port) were measured for 4 cm^{-1} and 8 cm^{-1}. One gold (reference) measurement of 512 scans was followed by eight sample measurements of 512 scans without moving the sample. These eight sample measurements were then averaged. A full measurement cycle took around 40 min, which included delays for purging the instrument after sample swap.

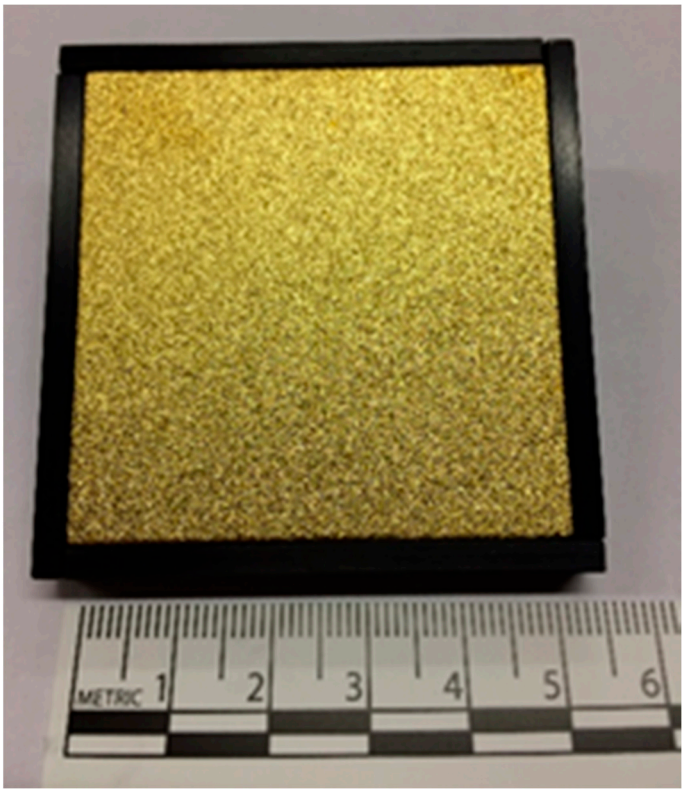

Figure A6. The Infragold reference standard used by UT-ITC for their calculation of reflectance.

Appendix E. Planetary Spectroscopy Laboratory (PSL) at the German Aerospace Center (DLR)

Currently PSL operates two identical Bruker VERTEX 80v vacuum FTIR spectrometer; one spectrometer is equipped with aluminium mirrors optimized for the UV, visible and near-IR spectral ranges, the second features gold-coated mirrors for the near to far IR spectral range. Hemispherical reflectance was measured under purging conditions, covering the 0.2 μm to above 200 μm spectral range. The two instruments share the collection of detectors, beamsplitters and optical accessories that are available in their equipment to cover a very wide spectral range, and this facilitates the cross-calibration procedures.

The instruments and the optical accessory units used are fully automatized and the data calibration and reduction are made with quality-controlled software, developed following the DLR quality management rules. Figure A7 shows the PSL laboratory. Two integrating spheres (one with gold coated surfaces, the other with PTFE coating) are available for hemispherical reflectance measurements. Reflectance measurements are calibrated by comparing with spectroscopic measurements of well characterized references (PTFE for UV, Spectralon for VIS and Infragold for MIR). Figure A8 shows the integrating sphere that was used to measure the two Optosol samples (1 and 2), the distilled water sample, and the reference (Infragold) that was used for calibration.

The measurement process is as follows. Firstly, a reference spectrum is acquired for each set-up. Then the spectrometer software allows measuring the sample and directly divides its spectrum for the measured reflectance. As a result, one has the relative reflectance of the sample to the reference used (Infragold in this case). By dividing for the real Infragold reflectance spectrum (provided from the producer of the reference itself), one gets the absolute hemispherical reflectance of the sample.

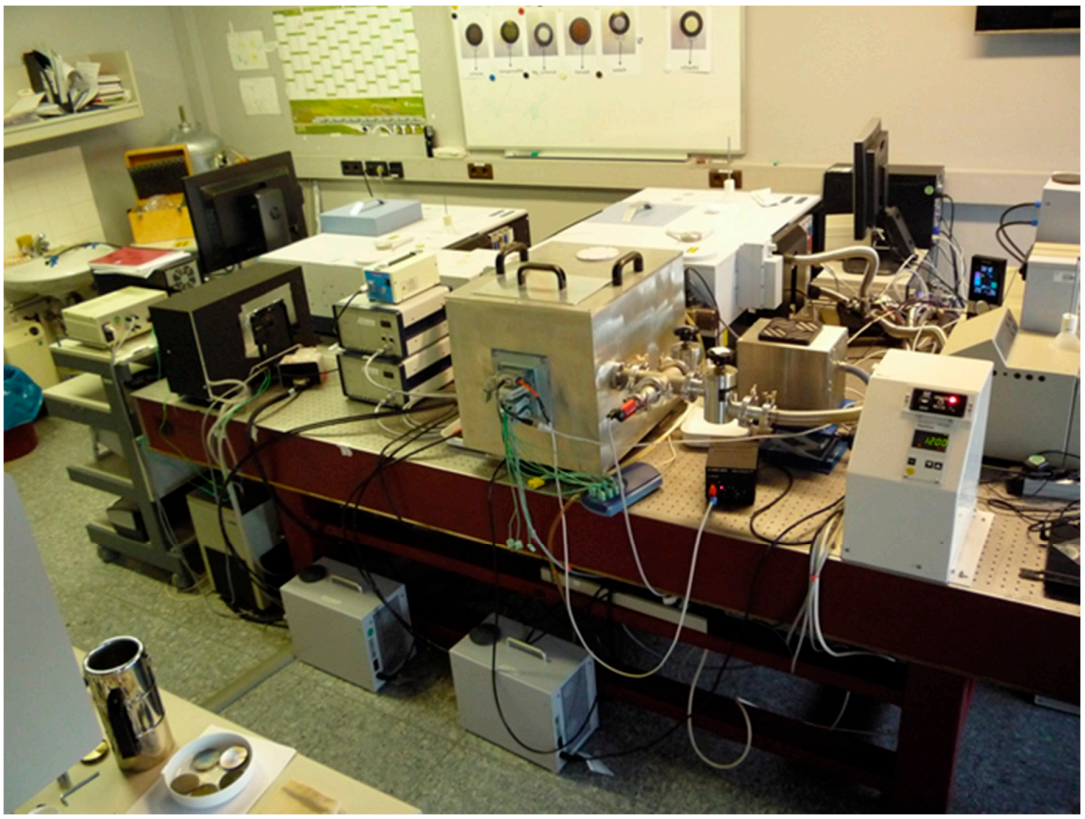

Figure A7. The laboratory setup at Planetary Spectroscopy Laboratory (PSL).

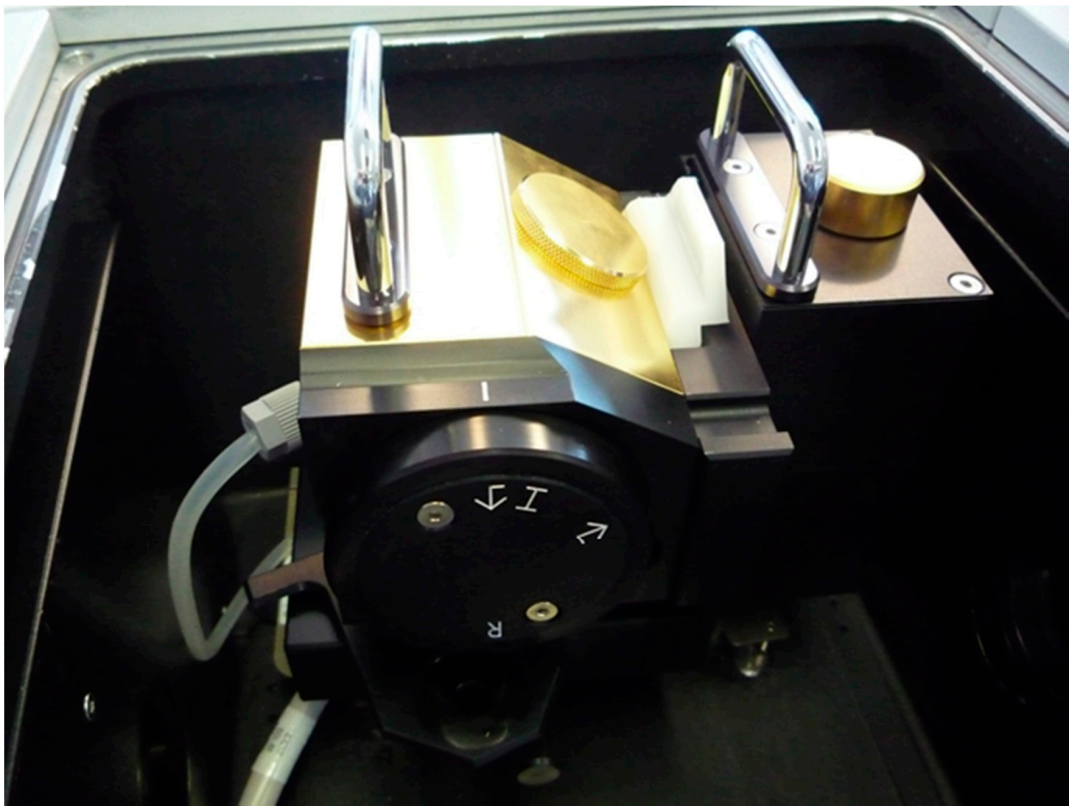

Figure A8. Gold coated integrating sphere and reference used for the experiment at PSL (German Aerospace Center (DLR)).

The number of scans depends on the spectral resolution. At 0.5 cm^{-1} resolution, 300 scans were collected for both the background and the sample, with a total running time from the start of background to the end of sample measurement of 30 min. At 4 cm^{-1} and 8 cm^{-1} resolution, 1000 scans were collected for both the background and the sample, with a total running time from the start of background to the end of sample measurement of 30 min and 16 min respectively.

Appendix F. ONERA—The French Aerospace Lab (ONERA)

The instrumentation is a Bruker Equinox 55 FTIR spectrometer equipped with a Labsphere 5inch (130 mm) diameter Infragold coated integrating sphere. The sample was placed at the bottom of the sphere, facing the 36mm diameter input port as shown in Figure A9. The sphere compartment was purged with nitrogen.

Figure A9. Bruker Equinox 55 with its integrating sphere in ONERA. The sample is on the lifting platform below the sphere.

The last Bruker conformity certificate was obtained on 22 November 2017, with a 2-year validity. The principal parameters of the set-up are summed-up in Table A4.

Table A4. Principal parameters used for the setup at ONERA.

Parameter	Setting
Spectral resolution	0.5 cm^{-1} (IRIS 2.1 mm), 4cm^{-1} (IRIS 6.1 mm) and 8 cm^{-1} (IRIS 7 mm)
Spectral range	2.5–1 6 µm (4000–625 cm^{-1})
Averaging	256 scans
TF parameters	Blackmann-Harris apodization, zero-filling factor 2

The measurement is a directional (13° incidence) hemispherical reflectance and obtained using four successive acquisitions of 256 scans each to compensate for the substitution error (see Figure A10). At 4 cm^{-1} resolution, a full measurement cycle took approximately 10 min.

Hemispherical reflectance denoted R_{dh} is then retrieved from these successive measurements using Equation (A3):

$$R_{dh} = \frac{M_{dir}^{ech}}{M_{dir}^{ref}} \times \frac{M_{indir}^{ref}}{M_{indir}^{ech}} \times R_{ref} \quad \text{(A3)}$$

where

M_{dir}^{ref}: Measurement of the standard sample (reference), direct beam;

M^{ech}_{ref}: Measurement of the sample, direct beam;
M^{ref}_{indir}: Measurement of the reference, tilted beam;
M^{ech}_{indir}: Measurement of the sample, tilted beam;
R_{ref}: the standard spectral reflectance (flat Labsphere IRS-94-020 Infragold sample; see Figure A11).

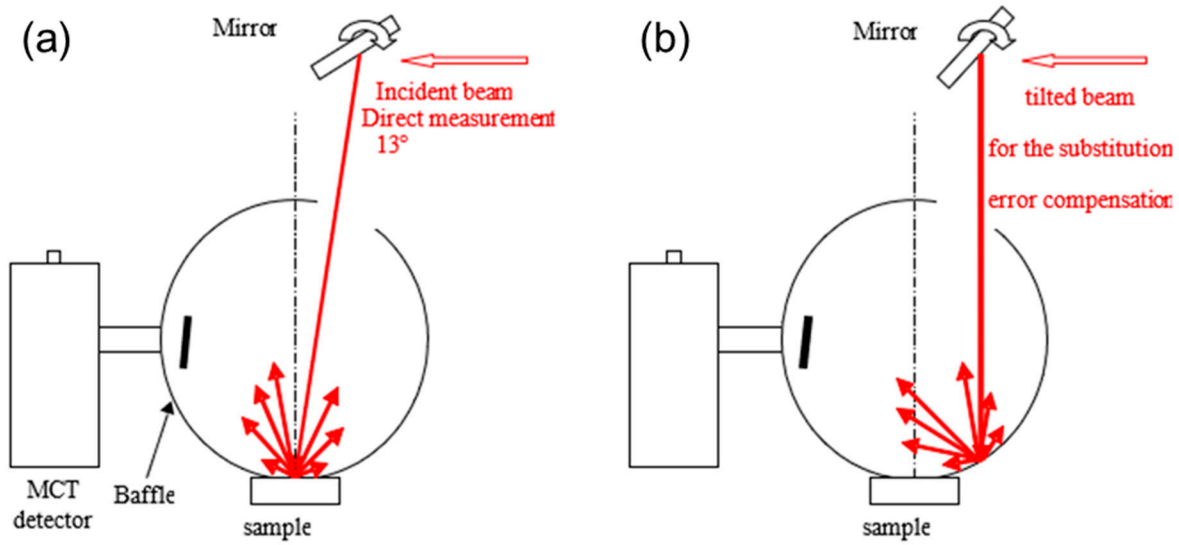

Figure A10. (**a**) Direct and (**b**) tilted beams for the compensation of the sphere substitution error at ONERA.

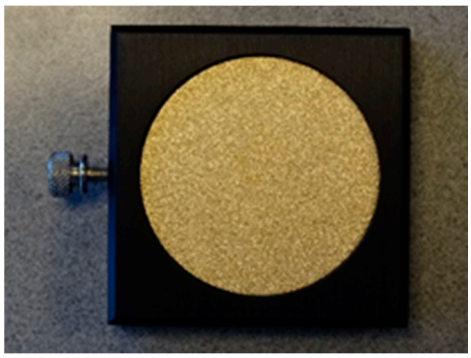

Figure A11. The Labsphere Infragold sample used at ONERA.

The measurement uncertainty can be approximated by Equation (A4):

$$\frac{\Delta R_{dh}}{R_{dh}} \approx \frac{\Delta\left[\frac{M^{ech}_{dir}}{M^{ref}_{dir}}\right]}{R_{dh}} + \Delta\left[\frac{M^{ref}_{indir}}{M^{ech}_{indir}}\right] + \frac{\Delta R_{ref}}{R_{ref}} \quad \text{(A4)}$$

The first term is both a bias and a noise contributor, the second one is a noise contributor while the last one is mainly a bias term. The first two terms are typically lower than 1–2%, outside the absorption bands and the last term depends on the standard reflectance knowledge given by Labsphere. Previous round robins and intercomparisons exhibit an overall error of approximately 3%.

Appendix G. Commonwealth Scientific and Industrial Research Organisation (CSIRO)

Measurements were performed at CSIRO using two setups. The first is a Bruker VERTEX 80v spectrometer with a Bruker integrating sphere attachment, mounted in the central sample chamber of the spectrometer, shown in Figure A12. The sphere is gold

coated with a speckled surface, with an entrance port on the right-hand side, a central mirror, which allows pointing of the beam on either the top or bottom ports. Both ports are knife edge, with the top port located on the upper right hemisphere of the sphere, whereas the bottom is directly below the mirror. The detector port is at the back, with 3 gold baffles. The bottom port is accessed by a sliding tray, which is sprung to move upwards to seal the sample to the port. This allows samples to only be approximately 50 mm × 50 mm in dimensions and approximately a maximum of 20 mm thick. Samples also must be flat. Round, aluminium cups are used to hold powders or soil samples, which are used in the bottom port. A gold plug is also provided, which is flat, resulting in the port having a flat surface. The top port has a curved plug, in contrast to the bottom port. There is potential that the reflectance from the top and bottom ports would produce different results due to the different path distance of the sample beam.

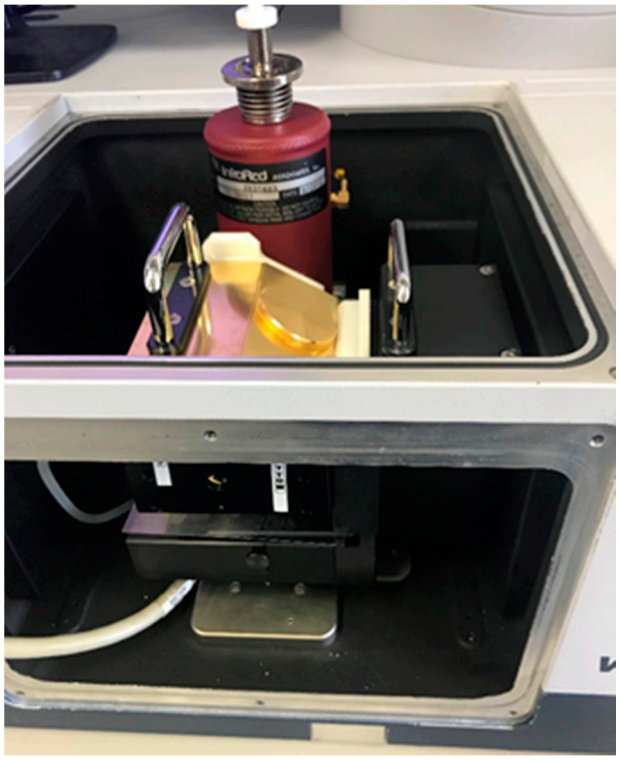

Figure A12. The Bruker integrating sphere and MCT detector in the Bruker VERTEX 80v sample compartment at CSIRO.

The sphere has a purge line built in, but it is unclear if it is operational. The upper sample port only allows for smaller samples, which preferably are flat. The Bruker VERTEX 80v has a sealed optics bench, allowing it to be purged and held under vacuum, reducing the interference of atmospheric gases.

The second setup was an Agilent 4300 Handheld FTIR, shown in Figure A13. The reference for this setup is a coarse silver target. The handheld data was only collected at 4 cm^{-1} resolution and only for a short 64 scan duration, which took approximately 60 s.

With the Bruker 80v spectrometer, measurements on Sample 1 and Sample 2 were collected with both the bottom port and top port for comparison. Due to the size of the laminate of the samples, it was required to hold them flat against the top port to clear the surrounding hardware (Teflon gasket between the entrance port) as shown in Figure A14.

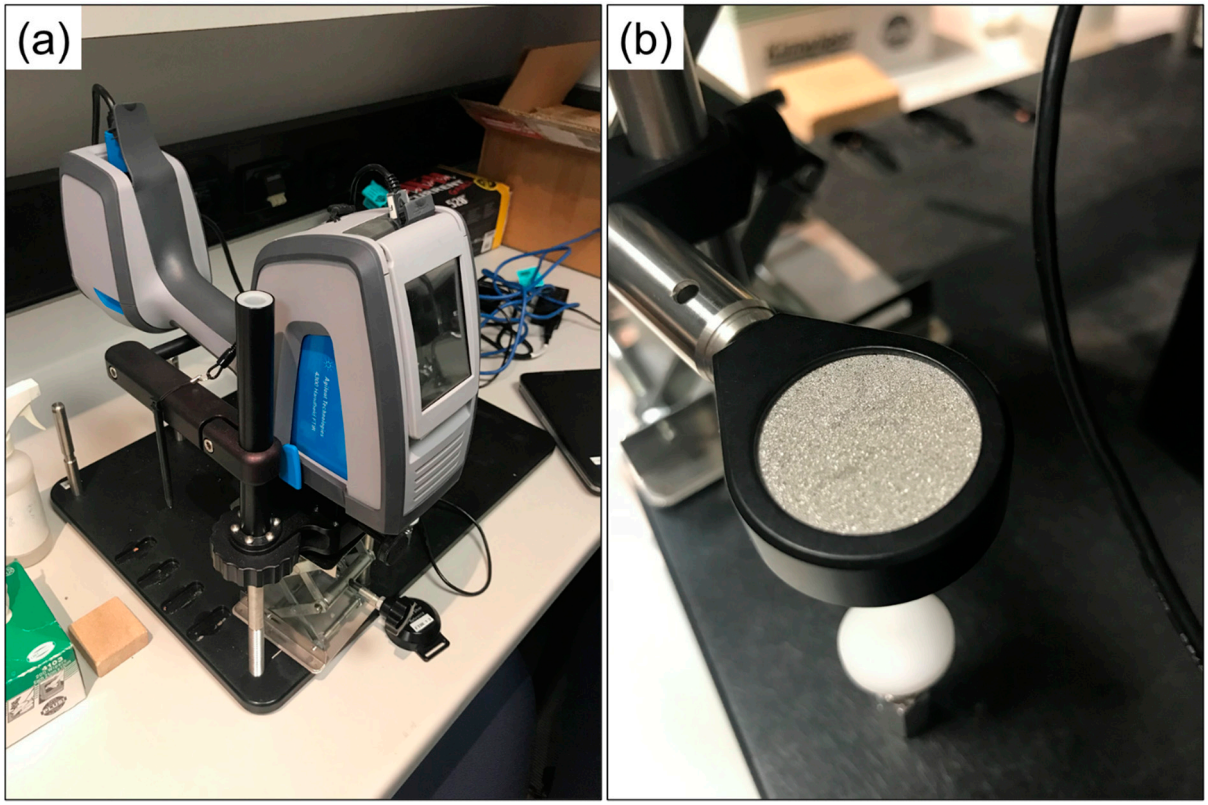

Figure A13. (**a**) The setup at CSIRO using the handheld Agilent 4300 FTIR with (**b**) the coarse silver reference target.

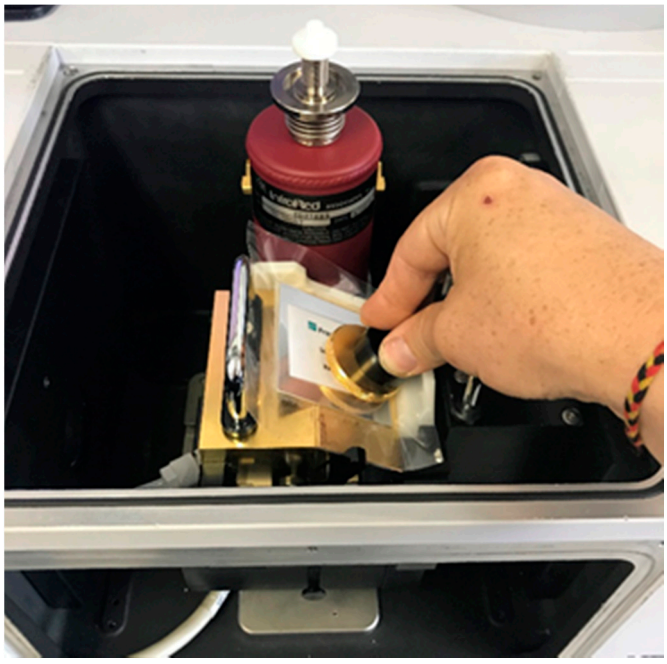

Figure A14. A view of Sample 1 being held in position over the top port.

A measurement was taken of the wall as a reference in both sample orientations as a comparison to the use of the opposite port. As such, measurements of each sample were made in four configurations (BG-B_S-T, BG-T_S-B, BG-W_S-T and BG-W_S-B where BG and S refer to reference and sample measurement positions respectively and B, T and W to the bottom port, top port or wall). The measurement of the wall was not exactly replicable each time, as a makeshift stopper was used to position the internal sphere mirror each time

(i.e., a pencil was placed in the level mechanism to approximate the wall location). In newer versions of this Bruker sphere a 3-position switch has been created by the manufacturer to allow for wall reference measurements, but the sphere at CSIRO is an older version, which does not have this capability. Parameters used for this experiment are shown in Table A5. For the Bruker FTIR, all sample and references measurements consisted of 256 scans. It took 110 s for the sample measurement and an additional 110 s for the background at 4 cm^{-1}, 48 s each at 8 cm^{-1} and 389 s each at 0.5 cm^{-1}.

The demineralised water was measured in the Bruker-supplied black aluminium sample cups and placed in the bottom port.

Table A5. Parameters used for the Bruker 80v FTIR spectrometer setup at CSIRO.

Parameter	Setting
Spectral resolution	4 cm^{-1} and 8 cm^{-1}
Spectral range	1.3–20 μm (7500–500 cm^{-1})
Averaging	256 scans
Aperture	6 mm
TF parameters	Blackmann-Harris apodization, zero-filling factor 2

References

1. Norman, J.M.; Becker, F. Terminology in thermal infrared remote sensing of natural surfaces. *Remote Sens. Rev.* **1995**, *12*, 159–173. [CrossRef]
2. World Meteorological Organization Essential Climate Variables. Available online: https://public.wmo.int/en/programmes/global-climate-observing-system/essential-climate-variables (accessed on 16 November 2020).
3. Li, Z.L.; Tang, B.H.; Wu, H.; Ren, H.; Yan, G.; Wan, Z.; Trigo, I.F.; Sobrino, J.A. Satellite-derived land surface temperature: Current status and perspectives. *Remote Sens. Environ.* **2013**, *131*, 14–37. [CrossRef]
4. Sobrino, J.A.; Mattar, C.; Pardo, P.; Jiménez-Muñoz, J.C.; Hook, S.J.; Baldridge, A.; Ibañez, R. Soil emissivity and reflectance spectra measurements. *Appl. Opt.* **2009**, *48*, 3664–3670. [CrossRef] [PubMed]
5. Jiménez-Muñoz, J.C.; Sobrino, J.A. A generalized single-channel method for retrieving land surface temperature from remote sensing data. *J. Geophys. Res. Atmos.* **2003**, *108*. [CrossRef]
6. Ermida, S.L.; Trigo, I.F.; Hulley, G.; DaCamara, C.C. A multi-sensor approach to retrieve emissivity angular dependence over desert regions. *Remote Sens. Environ.* **2020**, *237*, 111559. [CrossRef]
7. Pérez-Planells, L.; Valor, E.; Coll, C.; Niclòs, R. Comparison and Evaluation of the TES and ANEM Algorithms for Land Surface Temperature and Emissivity Separation over the Area of Valencia, Spain. *Remote Sens.* **2017**, *9*, 1251. [CrossRef]
8. Guillevic, P.; Göttsche, F.; Nickeson, J.; Hulley, G.; Ghent, D.; Yu, Y.; Trigo, I.; Hook, S.; Sobrino, J.A.; Remedios, J.; et al. Land Surface Temperature Product Validation Best Practice Protocol. Version 1.1. In *Best Practice for Satellite-Derived Land Product Validation*; CEOS LPV. 2018, p. 58. Available online: https://www.researchgate.net/publication/321537730_CEOS_WGCV_Land_Product_Validation_Subgroup_Land_Surface_Temperature_Product_Validation_Best_Practice_Protocol (accessed on 21 December 2020). [CrossRef]
9. Wan, Z.; Dozier, J. A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 892–905. [CrossRef]
10. Gillespie, A.; Rokugawa, S.; Matsunaga, T.; Steven Cothern, J.; Hook, S.; Kahle, A.B. A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (ASTER) images. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1113–1126. [CrossRef]
11. Trigo, I.F.; Peres, L.F.; Dacamara, C.C.; Freitas, S.C. Thermal Land Surface Emissivity Retrieved From SEVIRI/Meteosat. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 307. [CrossRef]
12. Sobrino, J.A.; Jiménez-Muñoz, J.C.; Sòria, G.; Gómez, M.; Ortiz, A.B.; Romaguera, M.; Zaragoza, M.; Julien, Y.; Cuenca, J.; Atitar, M.; et al. Thermal remote sensing in the framework of the SEN2FLEX project: Field measurements, airborne data and applications. *Int. J. Remote Sens.* **2008**, *29*, 4961–4991. [CrossRef]
13. Hulley, G.C.; Hook, S.J. Intercomparison of versions 4, 4.1 and 5 of the MODIS Land Surface Temperature and Emissivity products and validation with laboratory measurements of sand samples from the Namib desert, Namibia. *Remote Sens. Environ.* **2009**, *113*, 1313–1318. [CrossRef]
14. Sabol, D.E.; Gillespie, A.R.; Abbott, E.; Yamada, G. Field validation of the ASTER Temperature–Emissivity Separation algorithm. *Remote Sens. Environ.* **2009**, *113*, 2328–2344. [CrossRef]
15. Schmugge, T.; Ogawa, K. Validation of emissivity estimates from ASTER and MODIS data. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Denver, CO, USA, 31 July–4 August 2006; pp. 260–262.

16. Göttsche, F.; Olesen, F.-S.; Høyer, J.; Wimmer, W. *Fiducial Reference Measurements for Validation of Surface Temperature from Satellites (FRM4STS) Technical Report 3: A Framework to Verify the Field Performance of TIR FRM*; ESA: Noordwijk, The Netherlands, 2017.
17. Göttsche, F.-M.; Olesen, F.; Poutier, L.; Langlois, S.; Wimmer, W.; Santos, V.G.; Coll, C.; Niclos, R.; Arbelo, M.; Monchau, J.-P. *Report from the Field Inter-Comparison Experiment (FICE) for Land Surface Temperature*; ESA: Noordwijk, The Netherlands, 2018.
18. Salvaggio, C.; Miller, C.J. Comparison of field- and laboratory-collected midwave and longwave infrared emissivity spectra/data reduction techniques. In *Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII*; Shen, S.S., Descour, M.R., Eds.; SPIE: Orlando, FL, USA, 20 August 2001; Volume 4381, pp. 549–558. [CrossRef]
19. Hecker, C.A.; Smith, T.E.L.; Ribeiro Da Luz, B.; Wooster, M.J. Thermal infrared spectroscopy in the laboratory and field in support of land surface remote sensing. In *Thermal Infrared Remote Sensing: Sensors, Methods, Applications*; Kuenzer, C., Dech, S., Eds.; Springer: Dordrecht, The Netherlands, 2013; Volume 17, pp. 43–67.
20. Hulley, G.C.; Hook, S.J.; Abbott, E.; Malakar, N.; Islam, T.; Abrams, M. The ASTER Global Emissivity Dataset (ASTER GED): Mapping Earth's emissivity at 100 meter spatial scale. *Geophys. Res. Lett.* **2015**, *42*, 7966–7976. [CrossRef]
21. Meerdink, S.; Roberts, D.; Hulley, G.; Gader, P.; Pisek, J.; Adamson, K.; King, J.; Hook, S.J. Plant species' spectral emissivity and temperature using the hyperspectral thermal emission spectrometer (HyTES) sensor. *Remote Sens. Environ.* **2019**, *224*, 421–435. [CrossRef]
22. Mira, M.; Schmugge, T.J.; Valor, E.; Caselles, V.; Coll, C. Comparison of thermal infrared emissivities retrieved with the two-lid box and the TES methods with laboratory spectra. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1012–1021. [CrossRef]
23. Kotthaus, S.; Smith, T.E.L.; Wooster, M.J.; Grimmond, C.S.B. Derivation of an urban materials spectral library through emittance and reflectance spectroscopy. *ISPRS J. Photogramm. Remote Sens.* **2014**, *94*, 194–212. [CrossRef]
24. Nicodemus, F.E. Directional Reflectance and Emissivity of an Opaque Surface. *Appl. Opt.* **1965**, *4*, 767. [CrossRef]
25. Meerdink, S.K.; Hook, S.J.; Roberts, D.A.; Abbott, E.A. The ECOSTRESS spectral library version 1.0. *Remote Sens. Environ.* **2019**, *230*. [CrossRef]
26. Hulley, G.C.; Ghent, D.; Göttsche, F.M.; Guillevic, P.C.; Mildrexler, D.J.; Coll, C. Land Surface Temperature. In *Taking the Temperature of the Earth: Steps Towards Integrated Understanding of Variability and Change*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 57–127.
27. Hecker, C.; Hook, S.; van der Meijde, M.; Bakker, W.; van Werff, H.; Wilbrink, H.; van Ruitenbeek, F.; de Smeth, B.; van der Meer, F. Thermal infrared spectrometer for earth science remote sensing applications-instrument modifications and measurement procedures. *Sensors* **2011**, *11*, 10981–10999. [CrossRef]
28. Salisbury, J.W.; D'Aria, D.M. Emissivity of terrestrial materials in the 8-14 μm atmospheric window. *Remote Sens. Environ.* **1992**, *42*, 83–106. [CrossRef]
29. Hanssen, L.; Wilthan, B.; Filtz, J.-R.; Hameury, J.; Girard, F.; Battuello, M.; Ishii, J.; Hollandt, J.; Monte, C. Infrared spectral normal emittance/emissivity comparison. *Metrologia* **2016**, *53*, 03001. [CrossRef] [PubMed]
30. Hanssen, L.M.; Snail, K.A. Integrating Spheres for Mid-and Near-infrared Reflection Spectroscopy. In *Handbook of Vibrational Spectroscopy*; Chalmers, J.M., Griffiths, P.R., Eds.; Wiley: Chichester, UK, 2002; pp. 1–18.
31. Hardy, A.C.; Pineo, O.W. The Errors Due to The Finite Size of Holes and Sample in Integrating Spheres. *J. Opt. Soc. Am.* **1931**, *21*, 502. [CrossRef]
32. Jacquez, J.A.; Kuppenheim, H.F. Theory of the Integrating Sphere. *J. Opt. Soc. Am.* **1955**, *45*, 460. [CrossRef]
33. Goebel, D.G. Generalized Integrating-Sphere Theory. *Appl. Opt.* **1967**, *6*, 125. [CrossRef] [PubMed]
34. Gulmine, J.V.; Janissek, P.R.; Heise, H.M.; Akcelrud, L. Polyethylene characterization by FTIR. *Polym. Test.* **2002**, *21*, 557–563. [CrossRef]
35. Downing, H.D.; Williams, D. Optical constants of water in the infrared. *J. Geophys. Res.* **1975**, *80*, 1656–1661. [CrossRef]
36. Ullah, S.; Schlerf, M.; Skidmore, A.K.; Hecker, C. Identifying plant species using mid-wave infrared (2.5–6 μm) and thermal infrared (8-14μm) emissivity spectra. *Remote Sens. Environ.* **2012**, *118*, 95–102. [CrossRef]
37. Maturilli, A.; Helbert, J.; Arnold, G. The newly improved set-up at the Planetary Spectroscopy Laboratory (PSL). In *Infrared Remote Sensing and Instrumentation XXVII*; Strojnik, M., Arnold, G.E., Eds.; SPIE: San Diego, CA, USA, 9 September 2019; Volume 11128. [CrossRef]
38. Kato, S.; Matsunaga, T.; Tonooka, H. Statistical and in-situ validations of the ASTER spectral emissivity product at Railroad Valley, Nevada, USA. *Remote Sens. Environ.* **2014**, *145*, 81–92. [CrossRef]
39. Lesaignoux, A.; Fabre, S.; Briottet, X.; Olioso, A. Soil moisture impact on lab measured reflectance of bare soils in the optical domain [0.4–15 μM]. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Cape Town, South Africa, 12–17 July 2009; Volume 3.
40. Lau, I.C.; Ong, C.C.H.; Laukamp, C.; De Caritat, P.; Thomas, M. The acquisition and processing of voluminous spectral reflectance measurements of soils and powders for national datasets. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 4482–4484.
41. Wan, Z. A physics-based algorithm for retrieving land-surface emissivity and temperature from eos/modis data. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 980–996. [CrossRef]
42. Gupta, D.; Wang, L.; Hanssen, L.M.; Hsia, J.J.; Datla, R.U. *Standard Reference Materials: Polystyrene Films for Calibrating the Wavelength Scale of Infrared Spectrophotometers—SRM 1921*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 1995.

43. Davis, S.P.; Abrams, M.C.; Brault, J.W. *Fourier Transform Spectrometry*; Academic Press: San Diego, California, USA, 2001.
44. Bock, J.J.; Parikh, M.K.; Fischer, M.L.; Lange, A.E. Emissivity measurements of reflective surfaces at near-millimeter wavelengths. *Appl. Opt.* **1995**, *34*, 4812. [CrossRef]
45. Duan, S.B.; Li, Z.L.; Cheng, J.; Leng, P. Cross-satellite comparison of operational land surface temperature products derived from MODIS and ASTER data over bare soil surfaces. *ISPRS J. Photogramm. Remote Sens.* **2017**, *126*, 1–10. [CrossRef]
46. Malbéteau, Y.; Merlin, O.; Gascoin, S.; Gastellu, J.P.; Mattar, C.; Olivera-Guerra, L.; Khabba, S.; Jarlan, L. Normalizing land surface temperature data for elevation and illumination effects in mountainous areas: A case study using ASTER data over a steep-sided valley in Morocco. *Remote Sens. Environ.* **2017**, *189*, 25–39. [CrossRef]
47. Silvestri, M.; Romaniello, V.; Hook, S.; Musacchio, M.; Teggi, S.; Buongiorno, M.F. First comparisons of surface temperature estimations between ECOSTRESS, ASTER and landsat 8 over Italian volcanic and geothermal areas. *Remote Sens.* **2020**, *12*, 184. [CrossRef]
48. Berk, A.; Anderson, G.P.; Acharya, P.K.; Bernstein, L.S.; Muratov, L.; Lee, J.; Fox, M.; Adler-Golden, S.M.; Chetwynd, J.H.; Hoke, M.L.; et al. MODTRAN 5: A reformulated atmospheric band model with auxiliary species and practical multiple scattering options: Update. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*; Shen, S.S., Lewis, P.E., Eds.; SPIE: Honolulu, HI, USA, 2005; Volume 5806, p. 662.
49. Martins, J.P.; Coelho e Freitas, S.; Trigo, I.F.; Barroso, C.; Macedo, J. *Copernicus Global Land Operations-Lot I "Vegetation and Energy" Algorithm Theoretical Basis Document, Land Surface Temperature—LST V1:2, Issue I1.41*; Instituto Português do Mar e da Atmosfera: Lisbon, Portugal, 2019.
50. Tsilingiris, P.T. Comparative evaluation of the infrared transmission of polymer films. *Energy Convers. Manag.* **2003**. [CrossRef]
51. Sobrino, J.A.; Cuenca, J. Angular variation of thermal infrared emissivity for some natural surfaces from experimental measurements. *Appl. Opt.* **1999**, *38*, 3931. [CrossRef]
52. Hulley, G.C.; Duren, R.M.; Hopkins, F.M.; Hook, S.J.; Vance, N.; Guillevic, P.; Johnson, W.R.; Eng, B.T.; Mihaly, J.M.; Jovanovic, V.M.; et al. High spatial resolution imaging of methane and other trace gases with the airborne Hyperspectral Thermal Emission Spectrometer (HyTES). *Atmos. Meas. Tech.* **2016**, *9*, 2393–2408. [CrossRef]
53. Rousset-Rouviere, L.; Coudrain, C.; Fabre, S.; Ferrec, Y.; Poutier, L.; Viallefont, F.; Rivière, T.; Ceamanos, X.; Loke, T.; Fridman, A. SYSIPHE, an airborne hyperspectral imaging system from visible to thermal infrared. Results from the 2015 airborne campaign. In Proceedings of the 10th EARSEL SIG Imaging Spectroscopy, Zurich, Switzerland, 19–21 April 2017; pp. 1–15.
54. Vaughan, R.G.; Calvin, W.M.; Taranik, J.V. SEBASS hyperspectral thermal infrared data: Surface emissivity measurement and mineral mapping. *Remote Sens. Environ.* **2003**, *85*, 48–63. [CrossRef]
55. Göttsche, F.-M.; Olesen, F.-S.; Trigo, I.F.; Bork-Unkelbach, A.; Martin, M.A. Long Term Validation of Land Surface Temperature Retrieved from MSG/SEVIRI with Continuous in-Situ Measurements in Africa. *Remote Sens.* **2016**, *8*, 410. [CrossRef]
56. Sobrino, J.A.; Oltra-Carrió, R.; Jiménez-Muñoz, J.C.; Julien, Y.; Sòria, G.; Franch, B.; Mattar, C. Emissivity mapping over urban areas using a classification-based approach: Application to the Dual-use European Security IR Experiment (DESIREX). *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 141–147. [CrossRef]
57. Goddijn-Murphy, L.; Williamson, B. On Thermal Infrared Remote Sensing of Plastic Pollution in Natural Waters. *Remote Sens.* **2019**, *11*, 2159. [CrossRef]
58. Ermida, S.L.; Trigo, I.F.; DaCamara, C.C.; Göttsche, F.M.; Olesen, F.S.; Hulley, G. Validation of remotely sensed surface temperature over an oak woodland landscape—The problem of viewing and illumination geometries. *Remote Sens. Environ.* **2014**, *148*, 16–27. [CrossRef]